ARL 70-0278
DECEMBER 1970

AD718963

Aerospace Research Laboratories

MODERN TECHNIQUES IN ASTRODYNAMICS—AN INTRODUCTION

LYNN E. WOLAVER

APPLIED MATHEMATICS RESEARCH LABORATORY

D D C
FEB 26 1971
RECEIVED
B

PROJECT NO. 7071

This document has been approved for public release and sale; its distribution is unlimited.

AIR FORCE SYSTEMS COMMAND

United States Air Force

542

BLANK PAGE

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Agencies of the Department of Defense, qualified contractors and other government agencies may obtain copies from the

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

This document has been released to the

CLEARINGHOUSE
U. S. Department of Commerce
Springfield, Virginia 22151

for the sale to the public.

Copies of ARL Technical Documentary Reports should not be returned to Aerospace Research Laboratories unless return is required by security considerations, contractual obligations or notices on a specified document.

300 - January 1971 - CO305 - 23-71-388

ARL 70-0278

MODERN TECHNIQUES IN ASTRODYNAMICS – AN INTRODUCTION

LYNN E. WOLAVER

APPLIED MATHEMATICS RESEARCH LABORATORY

PROJECT NO. 7071

AEROSPACE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

FOREWORD

This report represents lecture notes of the author covering a course in Celestial Mechanics which has been given at the Air Force Institute of Technology for the past seven years. It should be noted that these are lecture notes and not a polished textbook. It is intended for engineers with a background in dynamics such as that covered in the books by Greenwood or Thompson. The author would also like to clearly state that these notes are by no means original with the author. The organization, misspellings and occasional jokes are those of the author. All else I have learned from others.

> "When 'Omer smote his bloomin' lyre,
> 'E'd heard men sing by land an' sea;
> An' what 'e thought 'e might require,
> 'E went an' took - the same as me!"
>
> -Kipling-Barracks-Room Ballads
> Introduction.

ABSTRACT

This report represents lecture notes for a graduate level course in celestial mechanics which has been given at the Air Force Institute of Technology. It covers a review of the two-body problem, discusses the three-body problem, the restricted three-body problem together with regularization and stability analysis. The main portion of the report develops the Hamilton-Jacobi theory and applies it to develop Lagrange's and Gauss' planetary equations. The oblate earth potential is developed and the secular equation solved. Effect of small thrust, drag, lunar-solar gravitational and solar radiation perturbations are developed mathematically and the net effects discussed. Von Zeipel's method for the solution of nonlinear equations is developed and used to solve Duffing's equation as an example. Special perturbations are discussed along with errors due to numerical integration and Encke's method is used to obtain approximate analytical results for the motion of stationary satellites. Finally a complete discussion of coordinate systems, time scale and astronomical constants are given. The report ends with a detailed discussion of the shape of the earth. Two appendices briefly cover numerical integration and a review of Lagrangian mechanics.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

BLANK PAGE

1. Introduction

Tycho Brake (1546-1601) spent 21 years peering at the sky and compiling a catalogue of the location of 777 stars and five of the planets. His precise observations were made entirely without the aid of a telescope. It was not until 1609 that Galileo first used the telescope for astronomical purposes. Tycho Brake's observatory and instruments are discussed in some detail in the Scientific American article by John Christianson [1]. From Tycho's data Johannes Kepler (1571-1630) was able to derive the three laws which bear his name. Kepler worked out a circular orbit for Mars that came within eight minutes of arc of Tycho's observations of that planet - a discrepancy amounting to no more than the thickness of a penny viewed edgewise at arm's length. But because Kepler knew that Tycho's data were accurate to within as little as four or five minutes of arc, he discarded the assumption of a circular orbit. He then tried fitting many curves, eventually arriving at an elliptical orbit that fitted the observations. Fortunately the observations were visual so that they did fit an ellipse. More accurate telescopic observations would have frustrated Kepler. From Keplers curve fitting came his three laws which he presented as unexplained facts. In 1609 in Astronomia Nova (De Motibus Stellae Martis) he wrote

(a) The heliocentric motions of the planets (i.e., their motion relative to the Sun) take place in fixed planes passing through the actual position of the Sun.

(b) The area of the sector traced by the radius vector from the Sun, between any two positions of a planet in its orbit, is proportional to the time occupied in passing from one position to the other.

(c) The form of a planetary orbit is an ellipse, of which the Sun occupies one focus.

Later in Harmonices Mundi (1619) he wrote

(d) The square of the periodic time is proportional to the cube of the mean distance (for an ellipse this is the semi-major axis).

Usually we call (c) the first law, and (d) Kepler's third law.

Almost as important as the laws themselves is the fact that perhaps for the first time in the history of science a theoretician had enough respect for the accuracy of a set of data that he abandon his original prejudices and searched for a totally new hypothesis.

The full dynamical implications of Kepler's laws were seen by Newton (1642-1727). Newton had formulated his laws of motion in 1687 in his Philosophial Naturalis Principia Mathematica (Principia) in three parts.

I. Every particle continues in its state of rest or uniform motion in a straight line unless it is acted upon by some exterior force.

II. The time rate of change of momentum of a particle is proportional to the force impressed upon it, and is in the direction in which the force is acting.

III. To every action there is an equal and oppositely directed reaction.

Given these fundamental laws and Kepler's observational data, Newton

was led to deduce the law of gravitational attraction. It is instructive to follow the arguments which led Newton to the conclusion. For a body to describe a plane curve, it is clear that the resultant of the forces acting on it must lie in the plane. Further, if a planet moving in its plane has polar coordinates r and f centered in the sun, the rate at which area A is swept out by the radius vector r is $dA/dt = 1/2\ r^2\dot{f}$. By Newton's second law of motion, the time derivative of $r^2\dot{f}$ is given by the moment, about the sun, of the resultant force acting on the planet, divided by the planet's mass. Hence, Kepler's second law (b) implied that the moment was zero, and therefore that the resultant force was central. This means its direction always passes through the sun. The constant $h = r^2\dot{f}$ is called the areal constant of the planet, and is equal to twice the rate at which area is described.

Kepler's "first" law (c) then implied that each planet was attracted to the sun by a force that varied like $1/r^2$. The argument is as follows:

The polar equation for an ellipse is

$$r = \frac{a\ (1-e^2)}{1+e\ \cos\ f} \qquad (1\text{-}1)$$

where a is the semi-major axis, e is the eccentricity (the distance from center to either focus divided by a), and f is measured from the pericenter (the point on an elliptical orbit closest to the occupied focus).

The time derivative of (1-1) gives

$$\dot{r} = \frac{ae\ (1-e^2)\ \dot{f}\ \sin\ \dot{f}}{(1+e\ \cos\ f)^2}\ , \qquad (1\text{-}2)$$

Using $\dot{f} = \frac{h}{r^2}$ this reduces to

$$\dot{r} = \frac{eh \sin f}{a\ (1-e^2)} \quad . \tag{1-3}$$

Similarly we have

$$\ddot{r} = \frac{eh \cos f}{a\ (1-e^2)}\ \dot{f} \qquad \text{or}$$

$$\ddot{r} = \frac{eh^2 \cos f}{ar^2(1-e^2)}$$

which can be written in the form

$$\ddot{r} = \frac{h^2}{r^3}\left[1 - \frac{r}{a\ (1-e^2)}\right] \quad . \tag{1-4}$$

The radial acceleration can now be calculated as follows:

$$a_r = \ddot{r} - r\dot{f}^2 = \frac{h^2}{r^3} - \frac{h^2}{r^3}\left(\frac{r}{a\ (1-e^2)}\right) - \frac{h^2}{r^3}$$

$$a_r = -\frac{h^2}{r^2 a\ (1-e^2)} \tag{1-5}$$

From this and Newton's second law we deduce that the resultant force acting on the planet is an attraction varying like $1/r^2$ since h, a, and e are all constants. Kepler's "third" law (d) enabled Newton to infer the complete mathematical form of the law of gravitation.

For any single planet, the area of the ellipse is $\pi a^2\sqrt{1 - e^2}$, and the rate of describing area is h/2. Thus the period P of the planet's orbital motion is

$$P = \frac{2\pi a^2\sqrt{1 - e^2}}{h} \tag{1-6}$$

The mean motion n in an elliptical orbit is defined as $2\pi/P$. Thus for any planet

$$n = \frac{h}{a^2\sqrt{1-e^2}} \tag{1-7}$$

$$n^2a^3 = \frac{h^2}{a(1-e^2)} \tag{1-8}$$

and thus

$$a_r = -\frac{h^2}{ar^2(1-e^2)} = -\frac{n^2a^3}{r^2} \tag{1-9}$$

By Kepler's "third" law, a^3 varies like P^2, so n^2a^3 is a constant that is the same for all planets. Thus the acceleration of each planet toward the sun is $\gamma m/r^2$. By Newton's third law, each planet exerts an attractive force on the sun varying like the mass of the planet; therefore, the gravitational attraction varies like the mass of the attracting body as well as like the mass of the attracted body. Thus, the law of gravitation was inferred.

"Every particle in the universe attracts every other particle with a force which is directly proportional to the product of the masses of the particles, and inversely proportional to the square of the distance between them."

$$F = -\frac{\gamma m_1 m_2}{r^2} \tag{1-10}$$

The mass in Newton's second law is inertial mass, i.e., mass times velocity = momentum. It occurred to Newton that the gravitation mass,

m_1 and m_2 might not be the same. He therefore experimented with pendulums of the same length and external shape, but filled with differing materials: gold, silver, glass, sand, salt, water, wheat, etc. These pendulums he found to have the same periods. To an accuracy of one part in 1000, he found the weight of a body near the earth's surface was always the same multiple of its inertial mass, regardless of the nature of its material. Thus, regardless of composition, the mass in the law of gravitation was the same parameter as the mass inertially defined. More recently this has been proven to be true to a constancy of about three parts in 10^{10}.

In 1665 when Newton first conceived the idea of universal gravity, he saw that the moon's motion around the earth ought to furnish a test. Since the moon's distance (as was known even then) is about 60 times the radius of the earth, the distance it should fall toward the earth in a second ought to be, if the law is correct, 1/3600 of 193 inches = 0.0535 inches (193 inches = the distance which a body falls in a second at the earth's surface). To see how much the moon is deflected from a straight line each second we have, according to the central force law, that if the moon's orbit was circular, its acceleration would be

$$a = \frac{4\pi^2 r}{t^2} \qquad (1\text{-}11)$$

and the deflection is one-half of this. If we compute using r = 238,840 miles reduced to inches and t, the number of seconds in a sidereal month, the deflection comes out 0.0534 inches, a difference of only 1/10,000 of an inch. Unfortunately, when Newton first made this test, the distance of the moon in miles was not known, because the size of the earth had not

then been determined with any accuracy. He used the length of a degree of latitude as 60 miles instead of 69 which is closer to the true value. Using this wrong earth radius, multiplying by 60, he obtained an earth-moon distance which was about 16 percent too small. He calculated a deflection of only 0.044 inches. This discordance was too great and he loyally abandoned the gravitational theory as being contradicted by fact!

Six years later, in 1671, Picard's measurement of an arc of a meridian in France corrected the error in the size of the earth and Newton, on hearing of it, at once repeated his calculations. The accordance was now satisfactory and he resumed the subject with zeal.

In Equation (1-10) the constant γ appears. This is sometimes called the Newtonian constant of gravitation and is not known with very great accuracy. It is approximately

$$\gamma = 6.670 \quad (1 \pm 0.0007) \times 10^{-8} \text{ cm}^3/\text{gm sec}^2 \tag{1-13}$$

This numerical constant depends on the units of mass, time, and distance.

Kepler's "third" law (d) did not take the action-reaction principle into account and to be precise we must include the effect of the mass of the planet. Thus, the correct form of Kepler's law is

$$\gamma(m_1 + m_2) = n^2 a^3 \tag{1-14}$$

with

$$n = \frac{2\pi}{P}.$$

Note that we had

$$a_r = \frac{n^2 a^3}{r^2} = \frac{\gamma(m_1 + m_2)}{r^2} \tag{1-15}$$

so that

$$n^2 a^3 = \gamma(m_1 + m_2). \tag{1-16}$$

This gives a simple relation between the units involved. Thus, for elliptical motion any one of the three unit quantities may be found when the other two are known. In most texts $\gamma = k^2$ and this is written

$$n^2 a^3 = k^2(m_s + m_e) \tag{1-17}$$

with k being called Gauss' constant. Gauss established its value by setting "a" equal to unity, taking the mass of the sun as unity, and m, the planet's mass, in units of solar mass. He used

$$m = \frac{1}{345710}$$

and took

$$n = \frac{2\pi}{P}$$

with P = 365.2563835

as the sidereal year in mean solar days.

This gives $k = \frac{2\pi}{\sqrt{1+m}\ P} = 0.01720209895000$

The two values on which it rests have been improved but to avoid making adjustments in k, it is regarded as absolutely accurate and the mean distance a of the earth's relative orbit is revised. More recent values have been more like

$$m = \frac{E}{S} = \frac{1}{332928} \qquad \frac{E+M}{S} = \frac{1}{328912}$$

$$T = 365.25635442.$$

This value of "a" so obtained from (1-17) is called the astronomical unit and is defined as that distance unit, such that if it is used, k will have exactly the value Gauss assigned to it. Even despite the uncertainties one should appreciate that k is known much more precisely than γ. This is a problem in units as some confusion exists if one writes

$$\gamma = k^2.$$

Thus γ is meant to refer to gram, cm., second units while k is the Gaussian constant of gravitation.

If we kept changing the value of k this would mean that all astronomical calculations would have to be repeated every time the value of m or the mass ratio was improved. Rather than do that, we accept Gauss' value of k as exact, we treat it as a definition, and let the value of a vary so as to make equation 17 hold true.

The result of this adjustment in a is that when more reliable data are obtained, all existing calculations have only to be scaled by a value of "a" defined by a slight variation of Kepler's law. This latter equation can be written as

$$n^2 a^3 = k^2 Ms(1 + \frac{Me}{Ms}) . \tag{1-18}$$

Using the values of n, and Me/Ms as determined by Gauss, this becomes

$$a^3 = 2.5226941 \times 10^{13} \ k^2 Ms \tag{1-19}$$

and this can be considered as defining the astronomical unit.

Note that it is often stated that the astronomical unit is the distance from the sun to a body of infinitesimal mass revolving in a circular orbit with a period of 365.24... ephemeris days, but this is a description only, not a definition.

Sometimes, a unit of time $1/k = 58.13244087$ days is chosen instead of the day; with such a unit $k^2 = 1$. Sometimes, as for double stars, the year is taken as the unit of time, the mean distance of the earth from the sun as the unit of distance, the solar mass as the unit of mass.

Then we have

$$(2\pi)^2 = k^2(1+m)$$

or

$$k = \frac{2\pi}{\sqrt{1+m}} \ . \tag{1-20}$$

There is another item worth noting and that is the careless usage of the words semi-major axis and the mean distance. Semi-major axis refers to an ellipse, the mean distance refers to the length in Kepler's equation. For the ideal two body case there is no problem; but when a third body of appreciable mass is introduced, Kepler's "third" law loses its geometrical significance. In such cases the law should be looked upon as what in fact it is, merely a definition of the unit of distance.

References:

Sterne, T. E., "An Introduction to Celestial Mechanics," Interscience, pp. 1-5 (1960).

Moulton, F. R., "An Introduction to Celestial Mechanics," Macmillan, pp. 29-35 (1914).

2. TWO BODY PROBLEM

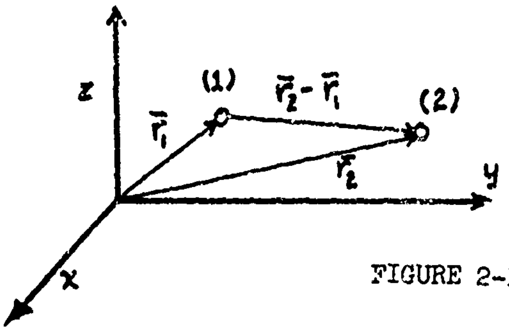


FIGURE 2-1

By Newton's law

$$m_i \frac{d^2\bar{r}_i}{dt^2} = k^2 \sum \frac{m_i m_j}{r_{ij}^3} (\bar{r}_j-\bar{r}_i), \ i = 1,2,\ldots n. \qquad (2\text{-}1)$$

For two bodies this becomes

$$m_1 \frac{d^2\bar{r}_1}{dt^2} = k^2 \frac{m_1 m_2}{r_{12}^3} (\bar{r}_2-\bar{r}_1)$$

$$m_2 \frac{d^2\bar{r}_2}{dt^2} = k^2 \frac{m_1 m_2}{r_{12}^3} (\bar{r}_1-\bar{r}_2) \qquad (2\text{-}2)$$

Subtracting gives

$$\frac{d^2}{dt^2} (\bar{r}_2 - \bar{r}_1) = -k^2 \frac{(m_1 + m_2)}{r_{12}^3} (\bar{r}_2 - \bar{r}_1) \qquad (2\text{-}3)$$

This is the equation of relative motion. When the mass of one of the bodies is very small compared to the other, i.e., when motion of m_2 has no effect on m_1, then with

$$\bar{r} = \bar{r}_2 - \bar{r}_1$$

$$r = r_{12}$$

$$\bar{v} = \frac{d\bar{r}}{dt}$$

$$\bar{R} = \frac{\bar{r}}{r} = \text{unit vector}$$

$$v = |\bar{v}|$$

we have

$$\frac{d^2\bar{r}}{dt^2} = \frac{d\bar{v}}{dt} = -k^2 \frac{(m_1 + m_2)}{r^2} \bar{R} = -f(r)\bar{R}. \tag{2-4}$$

Since $\bar{r}$ and $\bar{R}$ are collinear we have,

$$\bar{r} \times \frac{d\bar{v}}{dt} = 0 = \frac{d}{dt} (\bar{r} \times \bar{v}) \tag{2-5}$$

from which we find

$$\bar{r} \times \bar{v} = \bar{h} = \text{constant} = \text{angular momentum} \tag{2-6}$$

$\bar{r} \times \bar{v}$ is normal to both $\bar{r}$ and $\bar{v}$ and hence $\bar{r} \cdot \bar{h} = \bar{r} \cdot \bar{r} \times \bar{v} = 0.$

Thus motion remains in a fixed plane which is normal to $\bar{h}$. Now take the cross product of (2-4) times (2-6).

$$\frac{d\bar{v}}{dt} \times \bar{h} = -k^2 \frac{m_1 + m_2}{r^2} \bar{R} \times (\bar{r} \times \bar{v}) = k^2 (m_1 + m_2) \frac{d\bar{R}}{dt} \tag{2-7}$$

since

$$\bar{R} \times \left(\bar{R} \times \frac{d\bar{R}}{dt}\right) = \bar{R} \cdot \frac{d\bar{R}}{dt} \bar{R} - \bar{R} \cdot \bar{R} \frac{d\bar{R}}{dt} .$$

Thus

$$\frac{d}{dt}(\bar{v} \times \bar{h}) = k^2 (m_1 + m_2) \frac{d\bar{R}}{dt}. \tag{2-9}$$

This is called the Runge integral. Integrating gives

$$\bar{v} \times \bar{h} = k^2 (m_1 + m_2)(\bar{R} + \bar{e}) \tag{2-10}$$

$\bar{e}$ = arbitrary vector constant of integration.

We can find $\bar{e}$ as follows:

$$\bar{r} \cdot \bar{v} \times \bar{h} = k^2 (m_1 + m_2)\, \bar{r} \cdot (\bar{R} + \bar{e}) = k^2 (m_1 + m_2)(r + er \cos f) \tag{2-11}$$

(f is the angle between $\bar{r}$ and $\bar{e}$).

Noting that

$$\bar{r} \cdot \bar{v} \times \bar{h} = \bar{v} \cdot \bar{h} \times \bar{r} = \bar{h} \cdot \bar{r} \times \bar{v} = \bar{r} \times \bar{v} \cdot h = \bar{h} \cdot \bar{h} = h^2$$

one has

$$h^2 = k^2 (m_1 + m_2)(r + er \cos f). \tag{2-12}$$

This can be solved for r to give

$$r = \frac{h^2}{k^2 (m_1 + m_2)} \left[\frac{1}{1 + e \cos f} \right]. \tag{2-13}$$

Equation (2-13) can be interpreted geometrically by considering Figure (2-2).

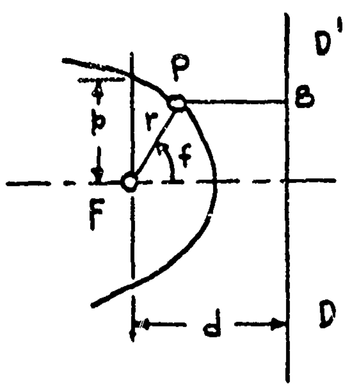


FIGURE (2-2) Conic Section.

Since a conic is the locus of a point (P) whose distance (r) from a given fixed point called the focus (F) is in a constant ratio (e) from a fixed line called the directix (DD'), one can write

$$r = e\,\overline{PB} = e(d - r\cos f)$$

which can be solved for r to give

$$r = \frac{ed}{1 + e\cos f} = \frac{p}{1 + e\cos f}$$

with $p \equiv ed = \text{constant}$

$\bar{e}$ is a certain vector perpendicular to $\bar{h}$ and $|e|$ = eccentricity.

The direction of $\bar{e}$ is along the major axis and f is measured from a line between focus and nearer vertex (perigee). Lets assume $0 \leq e < 1$, at $f = 0$

$$r_o = \frac{h^2}{k^2\,(m_1 + m_2)}\;\frac{1}{1 + e} \tag{2-14}$$

$f = \pi$

$$r_\pi = \frac{h^2}{k^2\,(m_1 + m_2)}\;\frac{1}{1 - e} \tag{2-15}$$

$$\begin{aligned} r_o + r_\pi = 2a &= \frac{h^2}{k^2\,(m_1 + m_2)}\left[\frac{1}{1 + e} + \frac{1}{1 - e}\right] \\ &= \frac{2h^2}{k^2\,(m_1 + m_2)\,(1 - e^2)} \end{aligned} \tag{2-16}$$

hence

$$r = \frac{a(1 - e^2)}{1 + e\cos f} \tag{2-17}$$

from which one can find

$$\cos f = \frac{a(1 - e^2)}{er} - \frac{1}{e}. \tag{2-18}$$

Note that at $f = 0$ we are at perigee and equation (2-17) gives,

$$r_p = a(1 - e)$$

while at $f = \pi$ we are at apogee and equation (2-17) reduces to

$$r_A = a(1 + e).$$

Note also that

$$\bar{r} \times \bar{v} = \bar{h} \tag{2-19}$$

with

$$|h| = |\bar{r}|\ |\bar{v}| \sin(\bar{r}, \bar{v}).$$

However since we have

$$|\bar{v}| \sin(\bar{r}, \bar{v}) = r \frac{df}{dt} \tag{2-20}$$

one can write

$$r^2 \frac{df}{dt} = h = 2 \frac{dA}{dt} \tag{2-21}$$

which is Kepler's "second" law.

In one period the radius sweeps out the entire area. The area of an ellipse is πab where b is the semi-minor axis and hence

$$\begin{aligned} 2\pi ab &= hP \\ P &= \frac{2\pi ab}{h} \end{aligned} \tag{2-22}$$

but

$$h^2 = a(1 - e^2)\, k^2\, (m_1 + m_2) \tag{2-23}$$

and (see Chapter 3),

$$a\ (1 - e^2) = \frac{b^2}{a}$$

so that

$$P = \frac{2\pi a^{3/2}}{k\sqrt{m_1 + m_2}} \tag{2-24}$$

this is the corrected form for Kepler's "third" law.

Returning to Equation (2-4) we can derive the vis-viva or energy integral.

$$\frac{d^2\bar{r}}{dt^2} = -k^2\ \frac{m_1 + m_2}{r^2}\ \bar{R} \tag{2-25}$$

$$\frac{d\bar{r}}{dt} \cdot \frac{d^2\bar{r}}{dt^2} = \frac{1}{2}\frac{d}{dt}\left(\frac{d\bar{r}}{dt} \cdot \frac{d\bar{r}}{dt}\right) = -k^2\ \frac{m_1 + m_2}{r^2}\frac{d\bar{r}}{dt} \cdot \bar{R}. \tag{2-26}$$

Since $\bar{r} = r\bar{R}$ then

$$\frac{d\bar{r}}{dt} = r\ \frac{d\bar{R}}{dt} + \bar{R}\ \frac{dr}{dt}$$

and hence

$$\frac{d\bar{r}}{dt} \cdot \bar{R} = r\ \frac{d\bar{R}}{dt} \cdot \bar{R} + \bar{R}\cdot\bar{R}\ \frac{dr}{dt} = \frac{dr}{dt}$$

so that we may write

$$\frac{1}{r^2}\frac{d\bar{r}}{dt} \cdot \bar{R} = \frac{1}{r^2}\frac{dr}{dt} = -\ \frac{d}{dt}\left(\frac{1}{r}\right). \tag{2-27}$$

Upon substitution of (2-27) into (2-26) we have

$$\frac{1}{2}\frac{d}{dt}\left(\frac{d\bar{r}}{dt} \cdot \frac{d\bar{r}}{dt}\right) = k^2\ (m_1 + m_2)\ \frac{d}{dt}\left(\frac{1}{r}\right).$$

This expression may be integrated to give

$$v^2 = k^2 \, (m_1 + m_2)\left(\frac{2}{r} + \text{constant}\right). \tag{2-28}$$

This must hold everywhere, in particular it must be true at perigee

$$v_p^2 = k^2 \, (m_1 + m_2)\left(\frac{2}{r_p} + C\right)$$

$$C = \frac{v_p^2}{k^2 \, (m_1 + m_2)} - \frac{2}{a \, (1 - e)} \tag{2-29}$$

where we have used $r_p = a \, (1 - e)$.

In addition, we can also write at perigee that

$$h = r_p v_p = a v_p \, (1 - e).$$

From Equation (2-23) we can write,

$$h^2 = a \, (1 - e^2) \, k^2 \, (m_1 + m_2) = r_p^2 v_p^2.$$

Solve for v_p to give

$$v_p^2 = \frac{a \, (1 - e^2) \, k^2 \, (m_1 + m_2)}{a^2 \, (1 - e)^2} .$$

When this is substituted for v_p in equation (2-29) we have

$$C = \frac{1 + e}{a \, (1 - e)} - \frac{2}{a \, (1 - e)} = - \frac{1}{a}$$

Equation (2-28) is then written as

$$\boxed{v^2 = k^2 \, (m_1 + m_2)\left[\frac{2}{r} - \frac{1}{a}\right]} \quad \text{VIS-VIVA} \tag{2-30}$$

Before proceeding further let's pause and review some properties of conic sections in general and of the ellipse in particular.

REFERENCES

1. Brand, L. Vector Analysis, John Wiley (1957).

2. Escobal, P.R., Methods of Orbit Determination, John Wiley (1965).

CHAPTER 3. CONIC SECTIONS

The conic sections are so named because they are the plane curves formed by the intersection of a plane and the surface of a right circular cone. The conic sections are shown in Figure 3-1.

The precise definition of a conic section is the locus of a point (P) which moves in a plane such that its distance from a fixed point called the focus (F) bears a constant ratio (e) to its distance from a fixed straight line called the directrix (DD').

$$e = \frac{r}{\ell} \qquad (3\text{-}1)$$

The relationship is shown in Figure 3-2 for the various conic sections. e is called the eccentricity. If

$e = 0$	the conic is a circle
$0 < e < 1$	the conic is an ellipse
$e = 1$	the conic is a parabola
$1 < e < \infty$	the conic is an hyperbola.

In each case the parameter a, which is half the maximum diameter, is called the semi-major axis of the conic section. Note that

$a = \infty$	for a parabola
$0 < a < \infty$	for an ellipse
$-\infty < a < 0$	for the hyperbolic case.

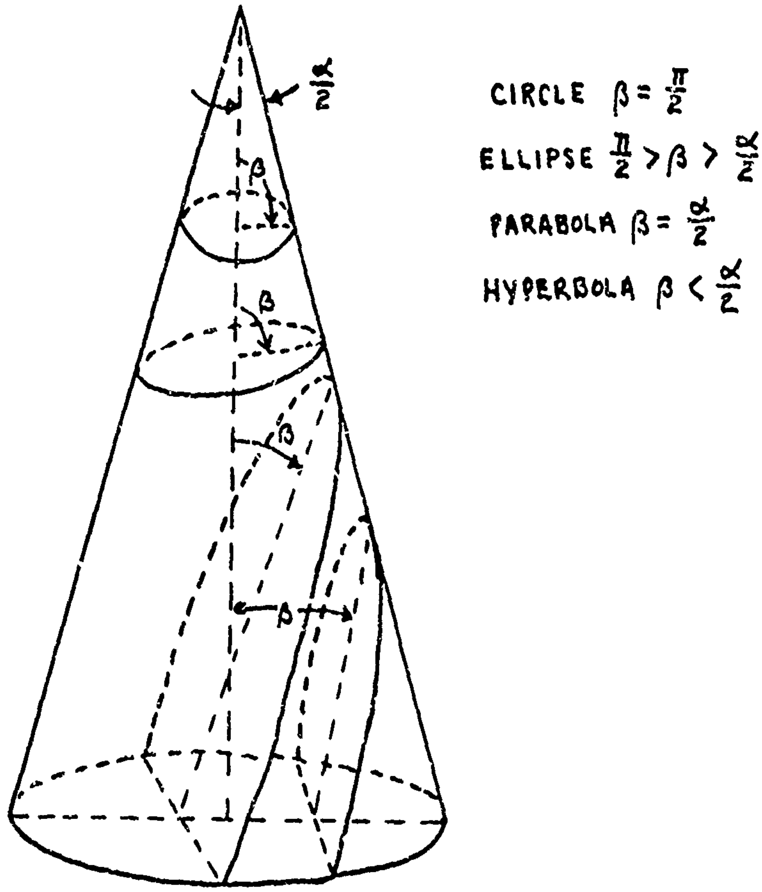


FIGURE 3-1 CONIC SECTIONS.

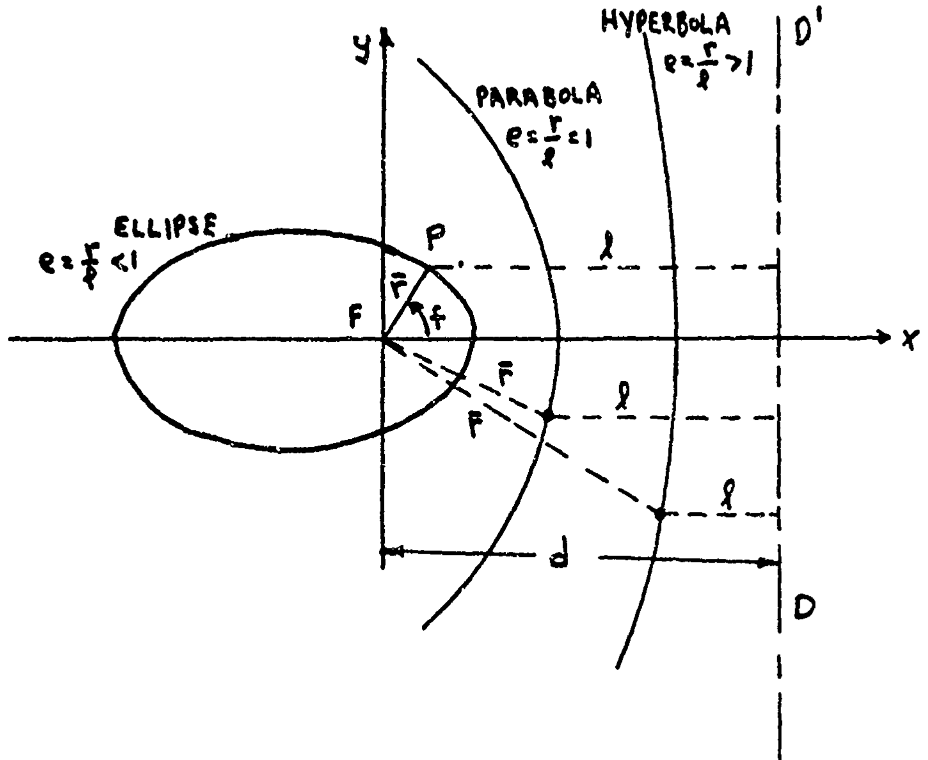


FIGURE 3-2 CONIC PARAMETERS

From Figure 3.2 and the definition of a conic section one can write

$$r = e \cdot \ell = e\,(d - r\cos f)$$

or (3.2)

$$r = \frac{ed}{1 + e\cos f}$$

The angle f is called the true anomaly. Some of the astronomical literature uses v for the true anomaly. Ballistic missile literature often uses θ and Kraft Ehricke, for one, uses η. When $f = 90^{\circ}$, equation (3-2) becomes

$$r = ed = p.$$

This defines the semi-latus rectum or conical parameter. When $f = 0$, we are at a point which is closest to the focus. This point is called the perifocus and for an orbit about the earth it is called the perigee; for an orbit about the sun it is called the perihelion. The point where the radius is a maximum occurs when $f = 180^{\circ}$ and is called the apofocus or apogee for the earth orbits and aphelion for sun orbits. From equation (3-2) we have the following

For $f = 0$, $r = r_p = \dfrac{p}{1 + e}$;

For $f = \pi$, $r = r_A = \dfrac{p}{1 - e}$.

For the ellipse the total largest diameter of the ellipse is denoted by 2a. From Figure 3-3 we conclude that

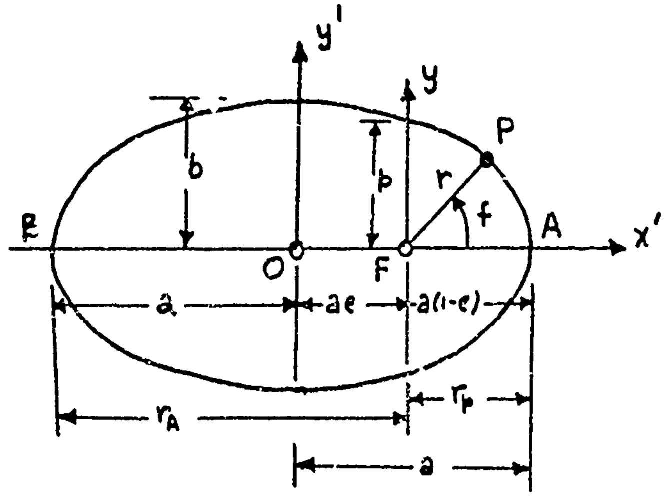


FIGURE 3-3 THE ELLIPSE.

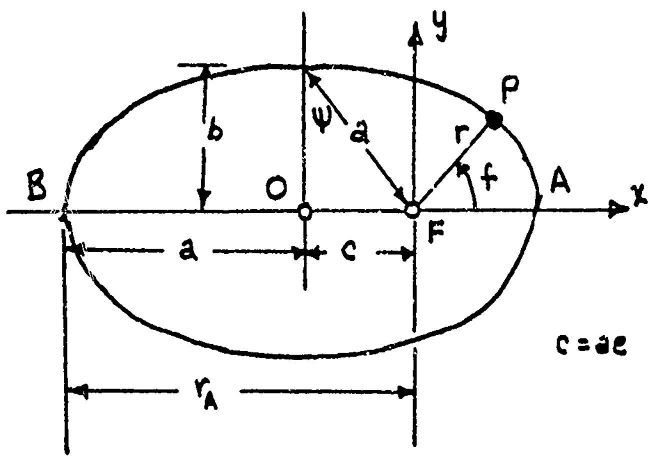


FIGURE 3-4

$$r_{max} + r_{min} = 2a$$

$$r_A + r_p = 2a \tag{3-3}$$

$$\frac{p}{1 - e} + \frac{p}{1 + e} = 2a$$

from which we find the relationship,

$$p = a(1 - e^2), \tag{3-4}$$

Equation (3-2) may thus be written as

$$r = \frac{a(1 - e^2)}{1 + e \cos f}\ . \tag{3-5}$$

Note that $r = r_p$ when $f = 0$, so from (3-5) we now have $r_p = a(1 - e)$ and since $r = r_A$ when $f = \pi$ we also have

$$r_A = a(1 + e). \tag{3-6}$$

This equation (3-6) may be written as

$$r_A = a + ae.$$

Referring to Figure 3-3 one sees that since OB = a, we must have OF = ae. This latter quantity is called the linear eccentricity, c.

$$c = ae$$

$$e = \frac{c}{a} \tag{3-7}$$

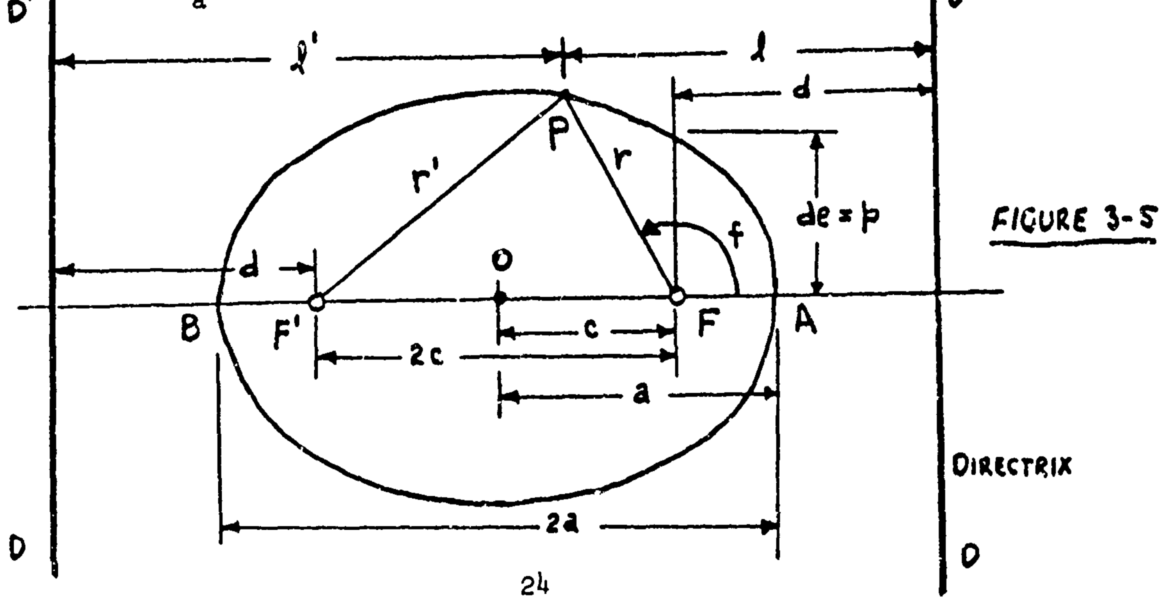


FIGURE 3-5

From Figure 3-5 we can show some important relationships. For any point P we have by symmetry that $r = e\ell$ and $r' = e\ell'$ and hence we can write

$$r + r' = e(\ell + \ell').$$

From Figure 3-5 we also have

$$\ell + \ell' = d + 2c + d = 2(d + c).$$

From this we can now write

$$r + r' = 2e(d + c). \tag{3-8}$$

At point A, the perigee, we can write

$$e = \frac{r}{\ell} = \frac{a - c}{d - (a - c)} \ .$$

This can be reduced to

$$d + c = \frac{a - c}{e} + a = \frac{a - c + ae}{e} = \frac{a}{e} \ . \tag{3-9}$$

Substitute (3-9) into (3-8) to give the relation,

$$r + r' = 2a \ . \tag{3-10}$$

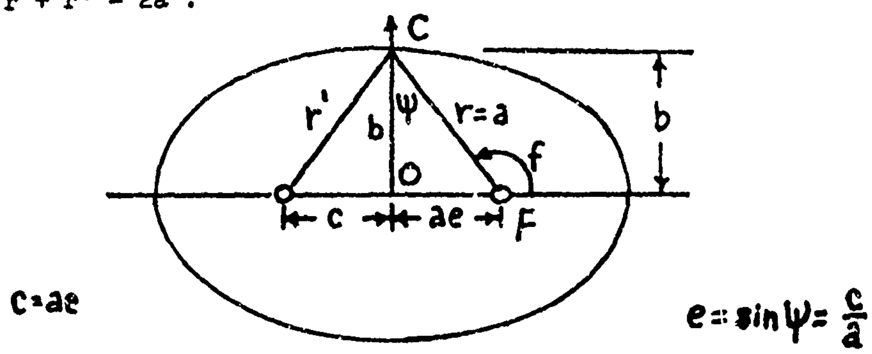


FIGURE 3-6

At the midpoint of the ellipse, Point C of Figure 3-6, we have $r = r'$ and the relation (3-10) gives

$$r = a.$$

By applying the Pythagorean theorem to the triangle of Figure 3-6 (ΔCOF) we have

$$a^2 = b^2 + c^2 = b^2 + a^2e^2.$$

Solve this for e^2 to obtain

$$e^2 = \frac{a^2 - b^2}{a^2} = 1 - \left(\frac{b}{a}\right)^2. \tag{3-11}$$

In addition since $p = a - ae^2$, we have

$$p = a - a\left(\frac{a^2 - b^2}{a^2}\right) \tag{3-12}$$

$$p = b^2/a.$$

Many other relationships may be similarly developed. Any reference book on Analytical Geometry will give many more handy relationships as for example that the area of an ellipse is πab, etc.

An important ancillary development uses the eccentric anomaly E.

To define the eccentric anomaly, E, one draws a circle of radius a about the center of the ellipse at O. (See Figure 3-7).

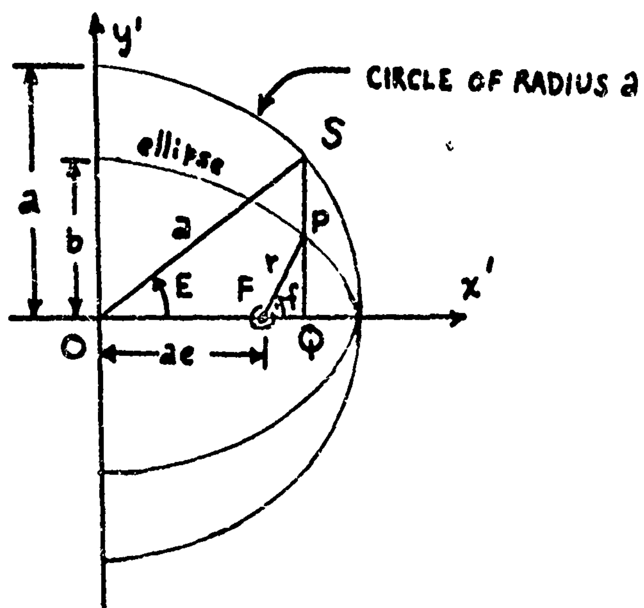


Figure 3-7 Eccentric Anomaly

The coordinate system is centered at the center of the ellipse at point 0. The x^1 coordinate, $\overline{0Q}$, of a point P on the ellipse is given by

$$\overline{0Q} = x^1 = a \cos E. \tag{3-13}$$

In this coordinate system the equation for an ellipse can be written as

$$\frac{(x^1)^2}{a^2} + \frac{(y^1)^2}{b^2} = 1. \tag{3-14}$$

Using equation (3-14) together with equation (3-13) one can obtain the ordinate value of the point P as,

$$\overline{PQ} = y^1 = b \sin E. \tag{3-15}$$

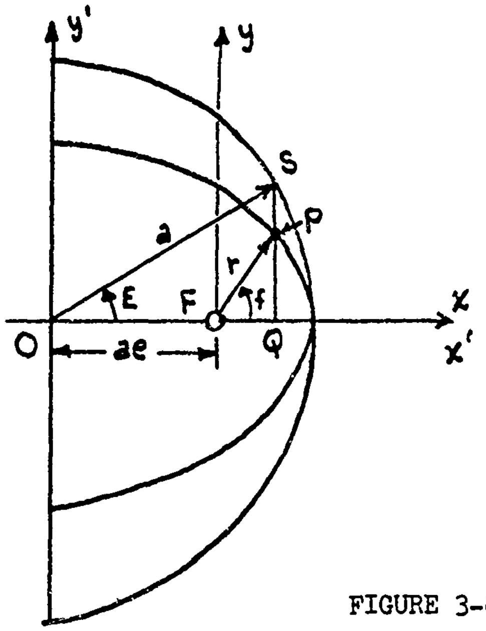


FIGURE 3-8

In a coordinate system centered at the focus(F) one has for a point P (from Figure 3-8),

$$x = \overline{FQ} = r \cos f = a \cos E - ae \tag{3-16}$$

and

$$y = \overline{PQ} = r \sin f = b \sin E = a\sqrt{1 - e^2} \sin E \tag{3-17}$$

This relation for b is found by solving equation (3-11).

From (3-16) and (3-17) one finds

$$r^2 = x^2 + y^2 = (a \cos E - ae)^2 + a^2(1 - e^2) \sin^2 E$$

$$r^2 = a^2(1 - e \cos E)^2 \qquad (3\text{-}18)$$

$$r = a(1 - e \cos E).$$

This relates the radius vector (drawn from the focus) with the eccentric anomaly.

Given

$$\left.\begin{aligned} x &= r \cos f = a(\cos E - e) \\ y &= r \sin f = a \sqrt{1 - e^2} \sin E, \end{aligned}\right\} \qquad (3\text{-}19)$$

the half angle formula,

$$\tan \frac{1}{2} f = \frac{\sin f}{1 + \cos f}$$

may be used to develop the fact that

$$\tan \frac{f}{2} = \sqrt{\frac{1 + e}{1 - e}} \left(\frac{\sin E}{1 + \cos E}\right) = \sqrt{\frac{1 + e}{1 - e}} \left(\tan \frac{E}{2}\right). \qquad (3\text{-}20)$$

The conic section can be represented by vectors as well. Let us define the vectors of interest as in Figure 3-9. From the basic definition of a conic section one can write

$$e = \frac{r}{\ell} = \frac{r}{d - \frac{\bar{d}}{d} \cdot \bar{r}}$$

$$r = de - \frac{e}{d} \bar{d} \cdot \bar{r}$$

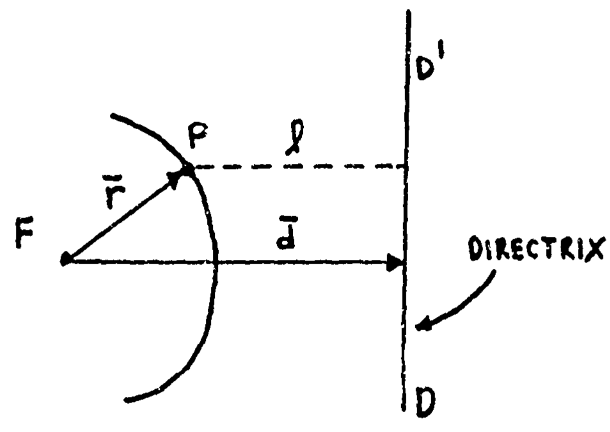
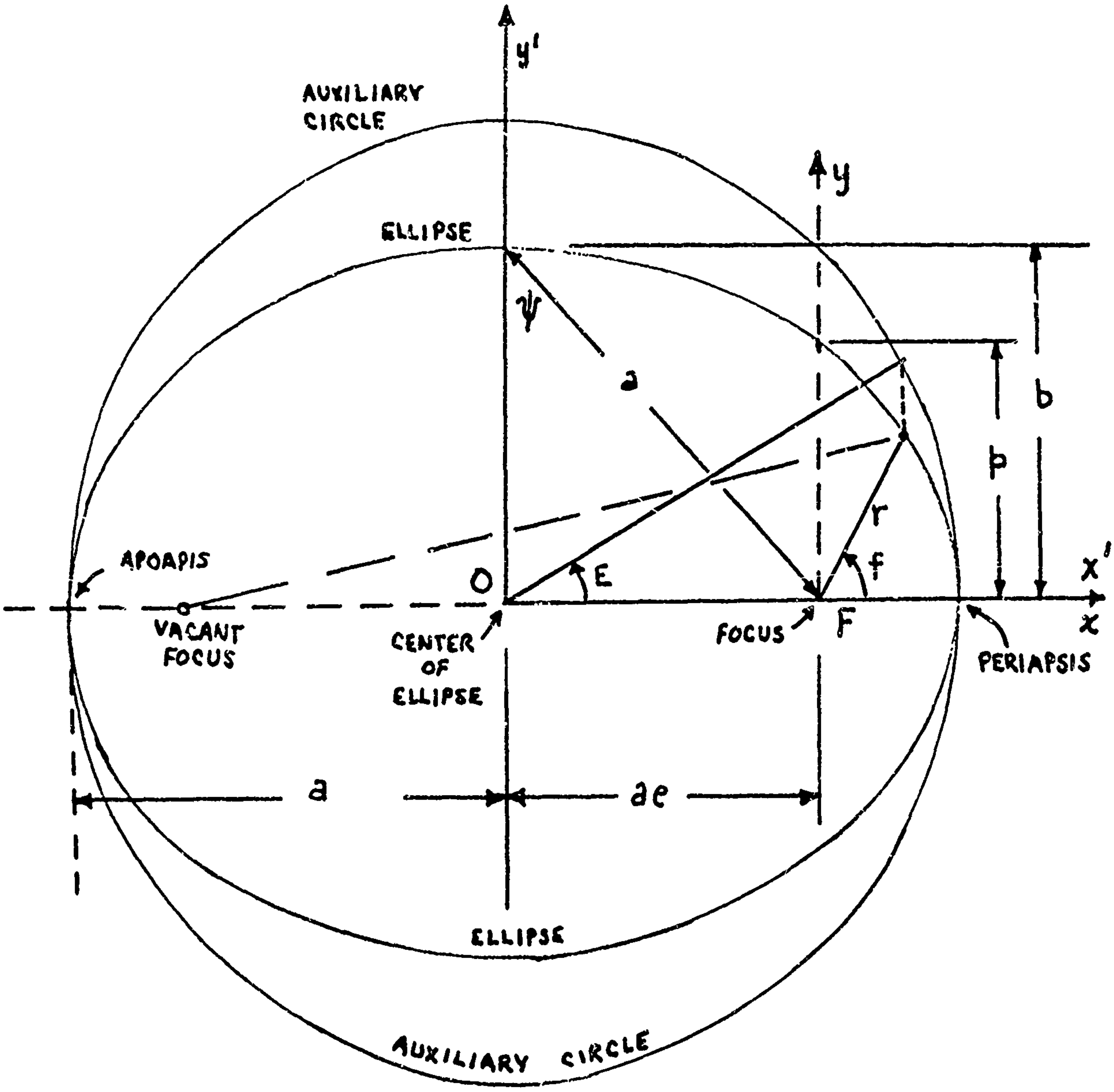


FIGURE 3-9 Vector Relations

REFERENCES

1. Taylor, Calculus with Analytical Geometry, Prentice-Hall, (1959).

2. Agnew, R.P., Analytic Geometry and Calculus, with Vectors, McGraw-Hill Book Company, (1962).



E = Eccentric anomaly

f = True anomaly

M = Mean anomaly

P = Period of orbit $n = \frac{2\pi}{P}$

$n(t-\tau_0) = M - e \sin E$

$a - r = ae \cos E$

ae = Linear eccentricity

e = Eccentricity = $\sqrt{1 - (b/a)^2} = \sin\Psi$

a = Semi-major axis

b = Semi-minor axis

p = Semi-latus rectum

$p = \frac{b^2}{a} = a(1-e^2)$

$b^2 = a^2(1-e^2)$

Area = $\pi ab = \pi a^2 \sqrt{1-e^2}$

$r = a(1-e \cos E)$

$\underbrace{\begin{bmatrix} x' = a \cos E \\ y' = b \sin E \end{bmatrix}}_{\left\{\begin{matrix}\text{origin}\\\text{at}\\\text{center}\end{matrix}\right\}}$ $\underbrace{\begin{bmatrix} x = r \cos f \\ y = r \sin f \end{bmatrix}}_{\left\{\begin{matrix}\text{origin}\\\text{at}\\\text{focus}\end{matrix}\right\}}$

FIGURE 3-10

CHAPTER 4. KEPLER'S EQUATION

Now let us return to equation (2-17) from which one can find equation (2-18) which is rewritten below.

$$f = \cos^{-1}\left(\frac{a(1 - e^2)}{er} - \frac{1}{e}\right). \tag{4-1}$$

We can differentiate this arc cosine relation to give

$$df = \frac{a\sqrt{1 - e^2}}{r\sqrt{a^2e^2 - (a - r)^2}}\ dr.$$

Using the relation $r^2 \frac{df}{dt} = h$, we can write

$$\frac{a\sqrt{1 - e^2}}{r\sqrt{a^2e^2 - (a - r)^2}}\ dr = \frac{h\ dt}{r^2}. \tag{4-2}$$

If we make use of the fact that

$$r = a - ae \cos E$$

we have

$$a - r = ae \cos E \tag{4-3}$$

and

$$dr = ae \sin E\ dE. \tag{4-4}$$

With these relations, equation (4-2) which can be written as

$$r \frac{dr}{\sqrt{a^2e^2 - (a - r)^2}} = \frac{hdt}{a\sqrt{1 - e^2}} \tag{4-5}$$

now becomes

$$(1 - e \cos E)\ dE = \frac{hdt}{a^2\sqrt{1 - e^2}} \tag{4-6}$$

which integrates to obtain

$$(E - e \sin E) = n (t - \tau_0) = M \tag{4-7}$$

with

$$n = \frac{h}{a^2\sqrt{1 - e^2}} = \frac{h}{ab} = \text{average angular velocity} \tag{4-8}$$

$$n = \frac{2\pi}{P} . \tag{4-9}$$

n is a constant called the mean motion. It is the angular rate in a circular orbit with the same period as the actual orbit. M is the mean anomaly. It is the angle which the radius would describe if it moved at a constant rate n. When $E = 0$, $M = 0$ then $t = \tau_0$; τ_0 is called the time of arrival at perigee, sometimes called the epoch.

Equation (4-7) is called Kepler's equation. There are 101 ways to solve it. We shall consider two. We have

$$M = n(t - \tau_0) = E - e \sin E. \tag{4-10}$$

This equation relates the time t with the eccentric anomaly E which in turn is related to r by means of the relation $r = a(1 - e \cos E)$. To find $r = r(t)$ we first pick t then solve (4-10) for E and thence find r for that instant of time. Recall also that

$$x = r \cos f = a(\cos E - e) \tag{4-11}$$

$$y = r \sin f = a\sqrt{1 - e^2} \sin E. \tag{4-12}$$

To solve Kepler's equation (4-10) we assume a series solution in powers of the eccentricity e, as

$$E = \sum_{k=1}^{\infty} C_k (M)e^k = E(e) \tag{4-13}$$

$$E(0) = M.$$

$$C_K(M) = \frac{1}{K!} \left. \frac{d^kE}{de^k} \right|_{e = 0}$$

If we write Kepler's equations as

$$F(E, e, M) = E - e \sin E - M = 0,$$

then using

$$\frac{dE}{de} = - \frac{\frac{\partial F}{\partial e}}{\frac{\partial F}{\partial E}}$$

we have

$$\frac{dE}{de} = \frac{\sin E}{1 - e \cos E} \text{ and hence } \frac{d^2E}{de^2} = \frac{\sin E \cos E}{(1-e \cos E)^2}, \text{ etc.} \tag{4-14}$$

We can now solve this differential equation by means of the power series to give terms like

$$E(o) = M \qquad E'(o) = \sin M, \text{ etc.}$$

to obtain

$$E = M + e \sin M + \frac{e^2}{2!} \sin 2M + \frac{e^3}{3!2^2} (3^2 \sin 3M - 3 \sin M) + \cdots \tag{4-15}$$

This becomes

$$\begin{aligned} E = M &+ \left(e - \frac{e^3}{8} + \frac{e^5}{192} - \frac{e^7}{9216} + \cdots\right) \sin M \\ &+ \left(\frac{e^2}{2} - \frac{e^4}{6} + \frac{e^6}{48} - \cdots\right) \sin 2M + \left(\frac{3e^3}{8} - \frac{27e^5}{128} + \frac{243e^7}{5120} + \cdots\right) \sin 3M \\ &+ \left(\frac{e^4}{3} - \frac{4e^6}{15} + \cdots\right) \sin 4M + \cdots \end{aligned} \tag{4-16}$$

This series converges for $e < 0.662743$. We can also use a differential correction technique. As a first approximation we take

$$E_o = M + e \sin M + \frac{e^2}{2} \sin 2M \tag{4-17}$$

then from

$$M = E - e \sin E$$

we find

$$\Delta M = M - M_o = \Delta E - e \cos E \, \Delta E$$

or

$$\Delta E_o = \frac{M - M_o}{1 - e \cos E_o} \quad . \tag{4-18}$$

The more correct value is then $E_1 = E_o + \Delta E_o$ from which we compute M_1 and get

$$\Delta E_1 = \frac{M - M_1}{1 - e \cos E_1} \text{ etc.} \tag{4.19}$$

This process converges very rapidly.

Similarly we can find a series expansion for f, the true anomaly,

$$\tan \frac{f}{2} = \sqrt{\frac{1 + e}{1 - e}} \tan \frac{E}{2} \ ,$$

thus given E we can find f. We obtain a series for f as a function of time ($M = n(t - \tau_o)$) similar to the one obtained for E to give the following:

$$f = M + \left(2e - \frac{e^3}{4} + \frac{5e^5}{96} + \frac{107e^7}{4608} + \cdot\cdot\right) \sin M$$
$$+ \left(\frac{5e^2}{4} - \frac{11e^4}{24} + \frac{17e^6}{192} + \cdot\cdot\right) \sin 2M + \left(\frac{13e^3}{12} - \frac{43e^5}{192} + \frac{95e^7}{512} - \cdot\cdot\cdot\right) \sin 3M$$
$$+ \left(\frac{103e^4}{96} - \frac{451e^6}{480} + \cdot\cdot\right) \sin 4M + \cdot\cdot\cdot\cdot . \tag{4-20}$$

These formulae and many more may be found in "Design Guide to Orbital Flight" McGraw-Hill Book Company. This same book can be ordered from DDC (ASTIA) as "Orbital Flight Manual," AD445453.

The two body problem can be summarized as follows:

Given a, e and τ_o (time of perigee passage), then

$$P = \frac{2\pi a^{3/2}}{k\sqrt{m_1 + m_2}} \qquad k = \text{Gauss's Constant} \tag{4-21}$$

$$n = \frac{2\pi}{P} \tag{4-22}$$

$$M = n\ (t - \tau_o) \tag{4-23}$$

$$E - e \sin E = n\ (t - \tau_o) = M. \tag{4-24}$$

Solve this for $E = E(t)$.

$$\begin{aligned} E = M &+ \left(e - \frac{e^3}{8} + \frac{e^5}{192} - \frac{e^7}{9216} + \cdots\right) \sin M \\ &+ \left(\frac{e^2}{2} - \frac{e^4}{6} + \frac{e^6}{48} + \cdots\right) \sin 2M + \left(\frac{3e^3}{8} - \frac{27e^5}{128} + \frac{243e^7}{5120} - \cdots\right) \sin 3M \\ &+ \left(\frac{e^4}{3} - \frac{4e^6}{15} + \cdots\right) \sin 4M + \left(\frac{125e^5}{384} - \frac{3125e^7}{9216} + \cdots\right) \sin 5M \\ &+ \left(\frac{27e^6}{80} - \cdots\right) \sin 6M + \left(\frac{16807e^7}{46{,}080} - \cdots\right) \sin 7M + \cdots. \end{aligned} \tag{4-25}$$

From this we find

$$r = a\ (1 - e \cos E) = \frac{a\ (1 - e^2)}{1 + e \cos f} \tag{4-26}$$

to give $r = r(t)$. We can also find $f = f(t)$ from the series expansion or from

$$\tan \frac{f}{2} = \sqrt{\frac{1 + e}{1 - e}}\ \tan \frac{E}{2} \tag{4-27}$$

For parabolic and hyperbolic orbits ($e \geq 1$) these formulae must be modified.

For a parabola, $e = 1$ and we have

$$r = \frac{p}{1 + \cos f} \quad . \tag{4-28}$$

Since the orbit is not periodic, there is no mean anomaly and Kepler's equation is not applicable. However since $h = r^2\dot{f}$ we can write

$$\int_{\tau_o}^{t} dt = \int_{o}^{f} \frac{r^2}{h}\, df = \frac{h^3}{\mu^2} \int_{o}^{f} \frac{df}{(1 + \cos f)^2} \quad . \tag{4-29}$$

$$\mu = k^2\,(m_1 + m_2).$$

Using the identities

$$1 + \cos f = 2 \cos^2 \frac{f}{2}$$

and

$$\sec^2 \frac{f}{2} = 1 + \tan^2 \frac{f}{2}$$

we have

$$t - \tau_o = \frac{h^3}{\mu^2} \int_{o}^{f} \sec^2 \frac{f}{2} \left(1 + \tan^2 \frac{f}{2}\right) df$$

$$t - \tau_o = 2 \sqrt{\frac{p^3}{\mu}} \left(\tan \frac{f}{2} + \frac{1}{3} \tan^3 \frac{f}{2}\right). \tag{4-30}$$

To find the position as a function of time one solves the cubic equation (4-30) for $\tan \frac{f}{2}$ and then finds f. We can then determine r from (4-28) or from the readily derived equation

$$r = \frac{1}{2}p \left(1 + \tan^2 \frac{f}{2}\right) \tag{4-31}$$

For the hyperbolic orbit the equivalent for equation (4-29) becomes

$$\int_{\tau_o}^{t} \frac{\sqrt{\mu}}{[-a\,(e^2 - 1)]^{3/2}}\, dt = \int_{o}^{f} \frac{df}{(1 + e \cos f)^2} \qquad (a < 0)$$

where we recall that for an hyperbola a is negative and $e > 1$. Integrate from τ_o to t and from 0 to f, using a table of integrals, to obtain

$$\sqrt{\frac{\mu}{-a^3}}\,(t - \tau_o) = e \sinh F - F \tag{4-32}$$

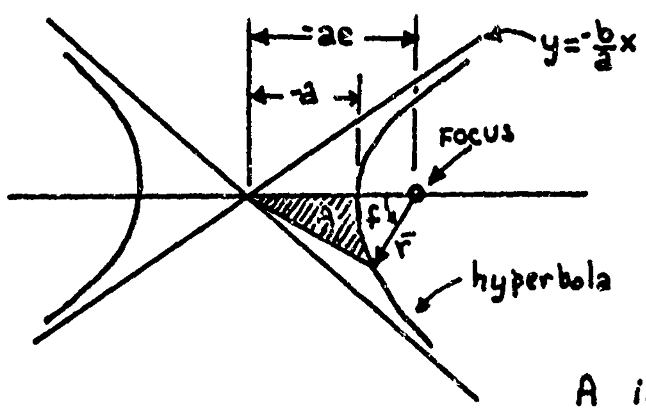
which is analogous to Kepler's equation. The radial equation can now be written as

$$r = a\,(1 - e \cosh F). \tag{4-33}$$

These hyperbolic equations are identical to elliptical motion under the substitution $E = -iF$. Thus one can write

$$\tan \frac{f}{2} = \sqrt{\frac{e+1}{e-1}}\,\tanh \frac{F}{2} \quad \text{etc.} \tag{4-34}$$

The inverse series solution can also be obtained. A good summary of these formulae is given in Kraft Ehricke's book, "Environment and Celestial Mechanics," Volume I of <u>Space Flight</u> starting on page 332. In addition see any of the handbooks referenced at the end of Chapter 5.



$$F = -\frac{A}{2\pi ab}$$

F is hyperbolic anomaly

A is area shown in figure.

CHAPTER 5. ELLIPTICAL PARAMETERS

In Chapter 3 we developed a few elliptical formulae such as $r_p = a(1 - e)$ and $r_A = a(1 + e)$. Using the fact that $r_A + r_p = 2a$ together with $r_p = a(1 - e)$ we have

$$e = 1 - \frac{r_p}{a} = \frac{r_A - r_p}{r_A + r_p} . \tag{5-1}$$

Similarly

$$b = a\sqrt{1 - e^2} = \sqrt{r_A r_p} , \tag{5-2}$$

$$\cos E = \frac{a - r}{ae} = \frac{a - r}{a - r_p} , \tag{5-3}$$

et cetera. These various interrelationships are summarized in the elliptical formulae table. The student should verify several of the relations. Another parameter of interest is the flight path angle. The angle γ is the inclination of the trajectory to the local horizontal.

Figure 5-1 below illustrates this.

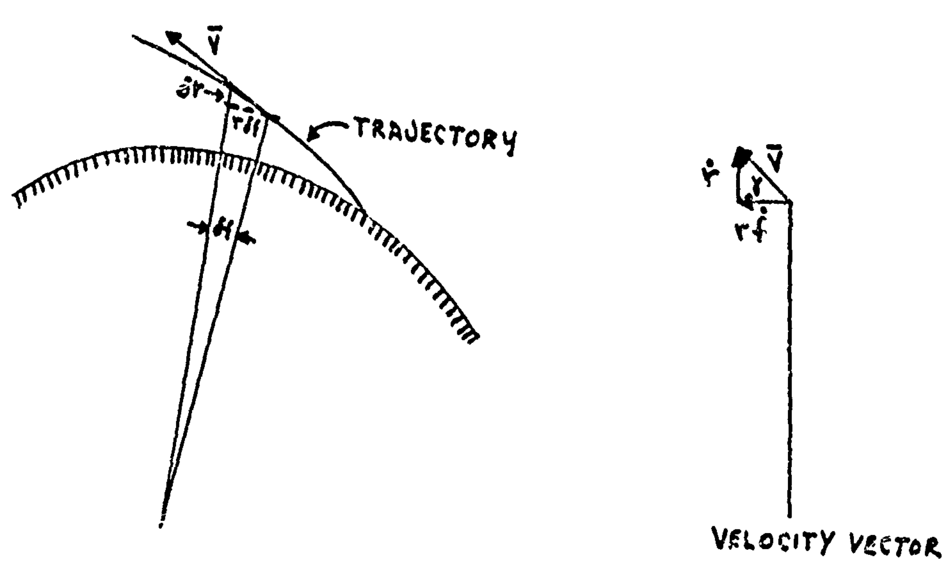


FIGURE 5-1

$$\tan \gamma = \lim_{\delta f \to 0} \frac{\delta r}{r \delta f} = \frac{1}{r} \frac{dr}{df}. \tag{5-4}$$

now

$$r = \frac{p}{1 + e \cos f}$$

therefore

$$\frac{dr}{df} = \frac{pe \sin f}{(1 + e \cos f)^2} \tag{5-5}$$

and hence

$$\tan \gamma = \frac{1 + e \cos f}{p} \frac{p (e \sin f)}{(1 + e \cos f)^2}$$

$$\tan \gamma = \frac{e \sin f}{1 + e \cos f} \quad . \tag{5-6}$$

The velocity at any point on the trajectory can be divided into its radial and tangential components,

$$\bar{v} = \dot{r} \, \bar{i} + r \, \dot{f} \, \bar{j} \ . \tag{5-7}$$

From Figure 5-1

$$\begin{aligned} r\dot{f} &= v \cos \gamma \\ \dot{r} &= v \sin \gamma. \end{aligned} \tag{5-8}$$

But recall $h = r^2\dot{f} = \text{constant}$, hence

$$rv \cos \gamma = h \tag{5-9}$$

This angular momentum, h, is conserved at every point of the ellipse. Energy is also conserved.

$$E = \text{kinetic} + \text{potential}$$

$$E = \frac{1}{2} v^2 - \frac{k^2 (m_1 + m_2)}{r} = \frac{1}{2} v^2 - \frac{\mu}{r} \tag{5-10}$$

where $\mu = k^2 (m_1 + m_2)$.

From the vis-viva

$$v^2 = \frac{2k^2\ (m_1 + m_2)}{r} - \frac{k^2\ (m_1 + m_2)}{a} \quad . \tag{5-12}$$

Substituting for v^2 in the energy expression gives,

$$E = -\frac{k^2\ (m_1 + m_2)}{2a} \tag{5-13}$$

Thus all elliptical trajectories with the same energy have the same semi-major axis length. This "a" value can be changed only by altering the total energy of the orbit. Equal energy orbits also have equal periods. Why?

Since the flight path angle, γ, is zero at apogee and perigee, we have

$$h = r_A v_A = r_p v_p \tag{5-14}$$

and since

$$r = \frac{h^2}{\mu\ (1 + e \cos f)}$$

then at apogee we have

$$r_A = \frac{r_p^2 v_p^2}{\mu\ (1 - e)} \quad . \tag{5-15}$$

Solve for e to give

$$e = 1 - \frac{r_p^2 v_p^2}{\mu r_A} = 1 - \frac{r_p}{r_A} \left(\frac{v_p}{v_c}\right)^2 \tag{5-16}$$

where $v_c^2 = \mu/r_p$. v_c is the velocity for a circular orbit at a distance r_p. Now substituting for e we have

$$r = \frac{h^2}{\mu\left[1 + \left(1 - \frac{r_p^2 v_p^2}{\mu r_A}\right)\cos f\right]} .$$

At perigee $r = r_p$, $f = 0$ giving

$$r_p = \frac{r_p^2 v_p^2}{\mu\left[1 + 1 - \frac{r_p^2 v_p^2}{\mu r_A}\right]}$$

Solving for r_A gives

$$r_A = \frac{r_p^2 v_p^2}{2\mu - r_p v_p^2}$$

or

$$\frac{r_A}{r_p} = \frac{\left(\frac{v_p}{v_c}\right)^2}{2 - \left(\frac{v_p}{v_c}\right)^2} \tag{5-17}$$

Put this in the equation for e (equation 5-16) to give,

$$e = \left(\frac{v_p}{v_c}\right)^2 - 1 . \tag{5-18}$$

v_p/v_c is the ratio of perigee velocity to the velocity needed to maintain a circular velocity at that radius (r_p). Thus for $v_p = v_c$, $e = 0$ and we have a circular orbit.

Having found the elliptical parameters in terms of the dynamic parameters (see page 55) we now have a large variety of two-parameter families, any one family of which will describe the ellipse. A further reduction can be made to the burn-out parameters.

The missile will reach some velocity v_b at a distance from the earth's center of r_b and at a flight path angle γ_b at the time of burn-out. We desire to find the elliptical parameters from these burn-out values. Since the total energy and angular momentum are conserved we have

$$E_b = \frac{1}{2} v_b^2 - \frac{\mu}{r_b} \quad ; \quad h_b = r_b v_b \cos \gamma_b ,$$

At perigee

$$E_p = E_b = \frac{1}{2} v_p^2 - \frac{\mu}{r_p} ; \quad h_p = h_b = r_p v_p .$$

Letting $v_s^2 = \mu / r_b$ the equivalent circular orbit velocity at the r_b altitude, we have the following development:

$$\frac{E_b}{v_s^2} = \frac{1}{2} \left(\frac{v_b}{v_s} \right)^2 - 1$$

$$\frac{\mu}{r_p} = - E_b + \frac{1}{2} v_p^2 \qquad v_p = \frac{r_b v_b \cos \gamma_b}{r_p}$$

$$\frac{\mu}{r_p} = \frac{r_b^2 v_b^2 \cos^2 \gamma_b}{2 r_p^2} - E_b$$

$$r_p^2 + r_p \left(\frac{\mu}{E_b} \right) - \frac{r_b^2 v_b^2 \cos^2 \gamma_b}{2 E_b} = 0$$

Solve for r_p to give

$$r_p = -\frac{\mu}{2E_b} \pm \frac{1}{2}\sqrt{\frac{\mu^2}{E_b^2} + \frac{2r_b^2 v_b^2 \cos^2 \gamma_b}{E_b}} \quad . \tag{5-19}$$

We use the - sign for r_p. One can show that the + sign corresponds to the apogee distance r_A. Now we have

$$\frac{E_b}{v_s^2} = \frac{1}{2}\left(\frac{v_b}{v_s}\right)^2 - 1$$

and since

$$\frac{v_b^2 r_b^2 E_b}{\mu^2} = \frac{v_b^2 E_b}{v_s^4} = \left(\frac{v_b}{v_s}\right)^2 \frac{E_b}{v_s^2}$$

we have

$$r_p = \frac{\mu}{v_s^2\left[2 - \left(\frac{v_b}{v_s}\right)^2\right]}\left[1 - \sqrt{1 + \left(\frac{v_b}{v_s}\right)^2 \cos^2 \gamma_b \left\{\left(\frac{v_b}{v_s}\right)^2 - 2\right\}}\,\right].$$

Since $\mu/v_s^2 = r_b$ we can write this as

$$\frac{r_p}{r_b} = \frac{1}{\left[2 - \left(\frac{v_b}{v_s}\right)^2\right]}\left[1 - \sqrt{1 + \left(\frac{v_b}{v_s}\right)^2\left[\left(\frac{v_b}{v_s}\right)^2 - 2\right] \cos^2 \gamma_b}\,\right].$$

Recall $v_p = \dfrac{v_b r_b \cos \gamma_b}{r_p}$

therefore

$$\frac{v_p}{v_b} = \frac{\cos \gamma_b}{\frac{r_p}{r_b}} \tag{5-21}$$

Hence given v_b, r_b and γ_b one can compute r_p and v_p. One additional item is of interest, the angular distance between burn-out and perigee.

$$\tan \gamma_b = \frac{e \sin f_b}{1 + e \cos f_b} .$$

Rearrangement gives

$$\sin f_b - \frac{1}{e} \tan \gamma_b = \tan \gamma_b \cos f_b.$$

Let $\nu = \sin f_b$ then $\cos f_b = \sqrt{1 - \nu^2}$.

Substitution and reduction gives

$$\nu^2 (1 + \tan^2 \gamma_b) - \frac{2\nu}{e} \tan \gamma_b + \frac{1 - e^2}{e^2} \tan^2 \gamma_b = 0.$$

Solving for ν we find

$$\sin f_b = \nu = \frac{\tan \gamma_b}{e (1 + \tan^2 \gamma_b)} \left[1 - \sqrt{1 - (1 - e^2)(1 + \tan^2 \gamma_b)} \right].$$

Now examine e. Recall that

$$h^2 = \mu a (1 - e^2).$$

Solve for e to give

$$e^2 = 1 - \frac{h^2}{\mu a} = 1 - \frac{r_b^2 v_b^2 \cos^2 \gamma}{\mu} \left[\frac{2}{r_b} - \frac{v_b^2}{u} \right]$$

where we have used the vis-viva evaluated at burn-out for the 1/a value, i.e.

$$\frac{-v_b^2}{\mu} + \frac{2}{r_b} = \frac{1}{a} .$$

This reduces to the following:

$$e^2 = 1 - \frac{r_b v_b^2 \cos^2 \gamma_b}{\frac{\mu}{r_b}} \left[\frac{2}{r_b} - \frac{v_b^2}{\frac{\mu}{r_b} r_b} \right]$$

$$e^2 = 1 - 2 \left(\frac{v_b}{v_s} \right)^2 \cos^2 \gamma_b + \left(\frac{v_b}{v_s} \right)^4 \cos^2 \gamma_b \, .$$

Hence

$$e = \left\{ 1 - \left[2 - \left(\frac{v_b}{v_s} \right)^2 \right] \left(\frac{v_b}{v_s} \right)^2 \cos^2 \gamma_b \right\}^{1/2} \tag{5-22}$$

We can then obtain e in terms of r_b, v_b and γ_b. From this we can then compute f_b in terms of the burn-out parameters.

If we let $c = \frac{v_p}{v_c}$ and $s = \frac{v_b}{v_s}$ with

$$v_c^2 = \frac{\mu}{r_p} \qquad v_s^2 = \frac{\mu}{r_b}$$

the formulae may be summarized as follows:

$$\frac{r_A}{r_p} = \frac{c^2}{2 - c^2} \tag{5-23}$$

$$e = c^2 - 1 = 1 - \frac{r_p}{r_A} c^2 \tag{5-24}$$

$$\frac{r_p}{r_b} = \frac{1}{2 - s^2} \left[1 - \sqrt{1 + s^2 (s^2 - 2) \cos^2 \gamma_b} \right] \tag{5-25}$$

$$\frac{r_A}{r_b} = \frac{1}{2 - s^2} \left[1 + \sqrt{1 + s^2 (s^2 - 2) \cos^2 \gamma_b} \right] \tag{5-26}$$

$$\frac{v_p}{v_b} = \frac{\cos \gamma_b}{\frac{r_p}{r_b}} \tag{5-27}$$

$$\sin f_b = \frac{\tan \gamma_b}{e\ (1 + \tan^2 \gamma_b)} \left[1 - \sqrt{1 - (1 - e^2)\ (1 + \tan^2 \gamma_b)}\right] \tag{5-28}$$

with

$$e = \sqrt{1 - (2 - s^2)\ s^2 \cos^2 \gamma_b} \tag{5-29}$$

These elements such as a, e and M or a, e and τ_o describe the ellipse in a given plane. We need to have three additional parameters to locate a particle in three dimensional space. To do this we adopt three Euler angles Ω, i, ω in order to locate the orbit plane and to orient the orbit within that plane. Consider Figure 5-2. We select the earth's equatorial plane as a reference plane (at least for earth satellites) and take the vernal equinox as the direction of the principal axis. The vernal equinox is the "point" where the sun crosses the equatorial plane of the earth. One should consult Chapter 32 for more details.

The angle measured in the equatorial plane between this principal axis and the line defining the intersection of the equatorial and orbit planes is called the longitude of the ascending node or nodal angle Ω. The angle between the orbit and equatorial planes is the inclination angle, i. The third angle is the argument of perifocus or the longitude of perigee (perihelion). This angle serves to locate the orientation of the major axis of the ellipse with respect to the line of nodes. The line of nodes is the line defined by the intersection of the equatorial and orbit planes. Refer to Figure 5-2.

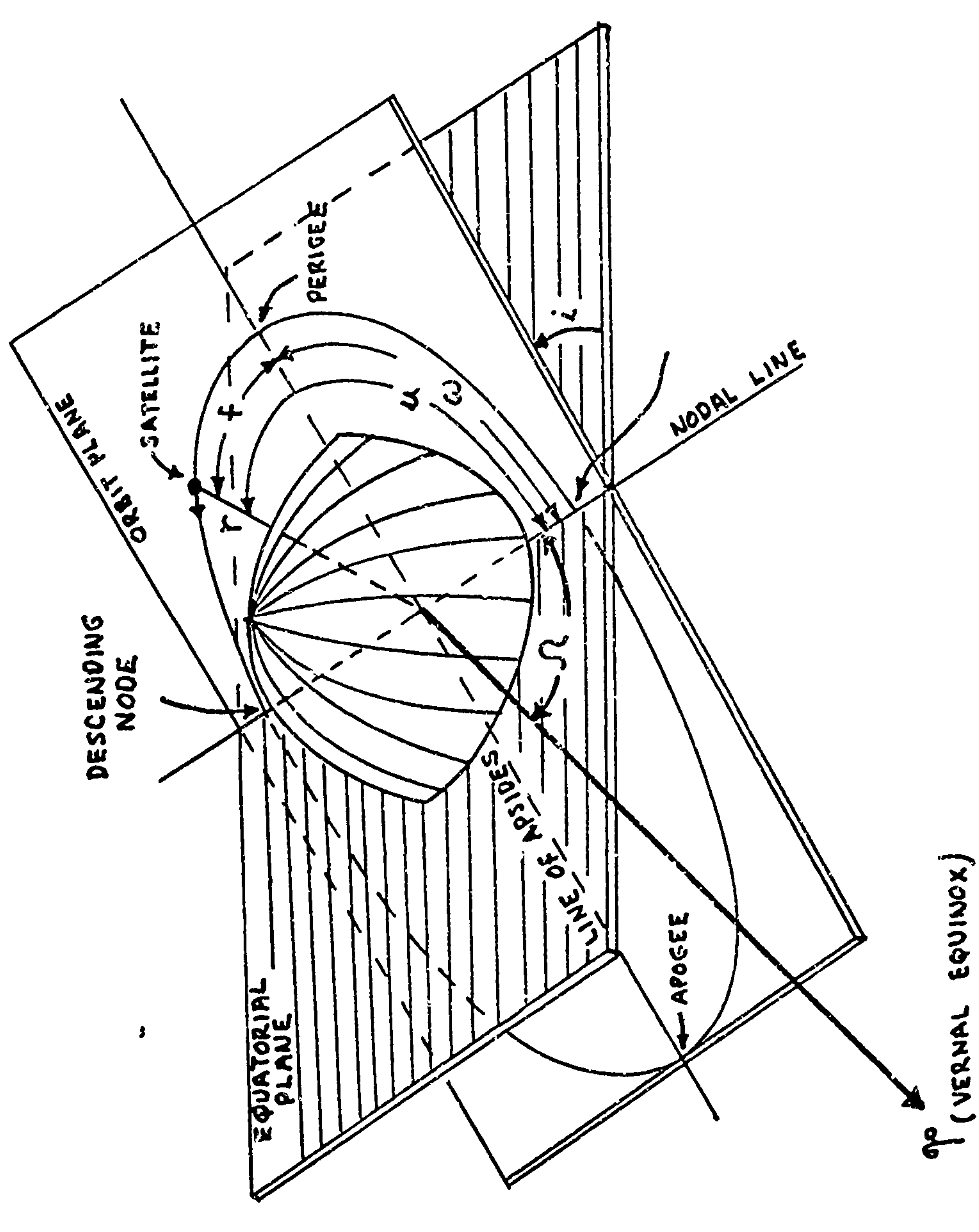


FIGURE 5-2 GEOCENTRIC ORBIT.

For three dimensions we choose the six elliptical elements.

Ω = longitude of ascending node $0 \leq \Omega \leq 2\pi$

i = inclination of orbital plane to reference plane $0 \leq i \leq \pi$

ω = argument of perigee $0 \leq \omega \leq 2\pi$

a = semi-major axis $0 < a < \infty$

e = eccentricity $0 \leq e < 1$

τ_o = time of perigee (perihelion) passage $0 \leq \tau_o < \infty$.

Sometimes $\tilde{\omega} = \pi = \Omega + \omega$ is used as it has geometrical significance.

Let us erect a rectangular coordinate system with its origin at the center of mass and choose the x-y plane to lie in the plane of reference with the x axis being the principal axis which points toward the vernal equinox. The z axis is perpendicular to the reference plane. How can one find the rectangular coordinates of a point P from the six elliptical elements? To develop this consider Figures 5-3 and 5-4.

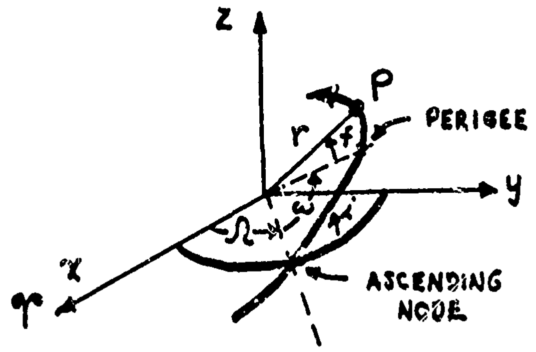


FIGURE 5-3

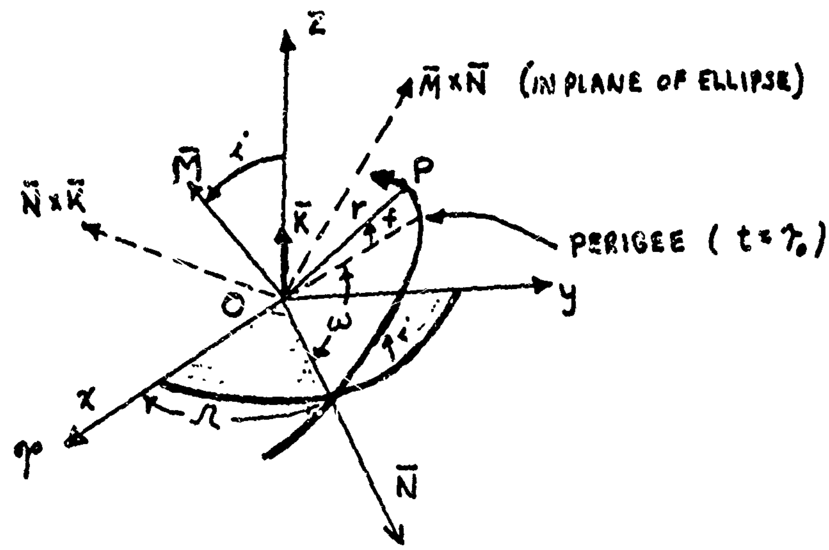


FIGURE 5-4

$\bar{M}$ is $\perp$ to plane of ellipse

$\bar{N}$ is along line of nodes

$\bar{I}$, $\bar{J}$, $\bar{K}$ are unit vectors along x, y, z

$\bar{M}$, $\bar{N}$, $\bar{M} \times \bar{N}$ are mutually $\perp$ vectors

$$\bar{r} = \overline{OP} = r \cos(\omega + f)\, \bar{N} + r \sin(\omega + f)\, \bar{M} \times \bar{N}$$

$$\bar{N} = \cos\Omega\, \bar{I} + \sin\Omega\, \bar{J}$$

$$\bar{M} = \cos i\, \bar{K} + \sin i\, \bar{N} \times \bar{K}$$

$$= \cos i\, \bar{K} + \sin i\, (-\cos\Omega\, \bar{J} + \sin\Omega\, \bar{I})$$

$$\bar{M} \times \bar{N} = \begin{bmatrix} \bar{I} & \bar{J} & \bar{K} \\ (\sin i \sin\Omega) & (-\sin i \cos\Omega) & (\cos i) \\ \cos\Omega & \sin\Omega & 0 \end{bmatrix}$$

so that

$$\bar{M} \times \bar{N} = \bar{I}\, (-\cos i \sin\Omega) + \bar{J}\, (\cos i \cos\Omega) + \bar{K} \sin i$$

$$\bar{r} = r \cos(\omega + f)\, \{\cos\Omega\, \bar{I} + \sin\Omega\, \bar{J}\}$$

$$+ r \sin(\omega + f)\, \{-\cos i \sin\Omega\, \bar{I} + \cos i \cos\Omega\, \bar{J} + \sin i\, \bar{K}\}$$

Finally

$$r = \bar{I}\,\{r\cos(\omega+f)\cos\Omega - r\sin(\omega+f)\cos i\sin\Omega\}$$
$$+ \bar{J}\,\{r\cos(\omega+f)\sin\Omega + r\sin(\omega+f)\cos i\cos\Omega\} \qquad (5\text{-}30)$$
$$+ \bar{K}\,\{r\sin(\omega+f)\sin i\}\ .$$

One thus obtains the relations:

$$x = r\,\{\cos(f+\omega)\cos\Omega - \sin(f+\omega)\cos i\sin\Omega\}$$
$$y = r\,\{\cos(f+\omega)\sin\Omega + \sin(f+\omega)\cos i\cos\Omega\} \qquad (5\text{-}31)$$
$$z = r\,\{\sin i\sin(f+\omega)\}\ .$$

One can also obtain relations for $\dot{x}$, $\dot{y}$ and $\dot{z}$ but these are best expressed as a function of the eccentric anomaly. If we define the direction cosines as

$$\ell_1 = \cos\Omega\cos\omega - \sin\Omega\sin\omega\cos i$$
$$m_1 = \sin\Omega\cos\omega + \cos\Omega\sin\omega\cos i \qquad (5\text{-}32)$$
$$n_1 = \sin\omega\sin i$$

then making use of the fact that

$$r\cos f = a\,(\cos E - e)$$

$$r\sin f = a\sqrt{1-e^2}\ \sin E = b\sin E,$$

equation (5-31) can be written as

$$x = a\ell_1\cos E + b\ell_2\sin E - ae\ell_1$$
$$y = am_1\cos E + bm_2\sin E - aem_1 \qquad (5\text{-}33)$$
$$z = an_1\cos E + bn_2\sin E - aen_1$$

where

$$\ell_2 = -\cos\Omega\sin\omega - \sin\Omega\cos\omega\cos i$$
$$m_2 = -\sin\Omega\sin\omega + \cos\Omega\cos\omega\cos i \qquad (5\text{-}34)$$
$$n_2 = \cos\omega\sin i.$$

The only quantity which is a function of time on the right hand side is E, the eccentric anomaly given by

$$n(t - \tau_o) = E - e \sin E. \tag{5-35}$$

Take the derivative with respect to time to give

$$\dot{E} = \frac{dE}{dt} = \frac{n}{1 - e \cos E} = \frac{na}{r} . \tag{5-36}$$

The equation (5-33) may be differentiated to give terms like

$$\dot{x} = -a\ell_1 \sin E \frac{dE}{dt} + b\ell_2 \cos E \frac{dE}{dt} \text{ etc.} \tag{5-36}$$

Substitution of $\dot{E}$ reduces these to the set

$$\begin{aligned} \dot{x} &= \frac{na}{r} (b\ell_2 \cos E - a\ell_1 \sin E) \\ \dot{y} &= \frac{na}{r} (bm_2 \cos E - am_1 \sin E) \\ \dot{z} &= \frac{na}{r} (bn_2 \cos E - an_1 \sin E). \end{aligned} \tag{5-37}$$

For the reverse problem, let us find the elliptical elements in terms of the position and velocity at a given time. Recall

$$\begin{aligned} \bar{r} &= \bar{i}x + \bar{j}y + \bar{k}z \\ \bar{v} &= \bar{i}\dot{x} + \bar{j}\dot{y} + \bar{k}\dot{z} \quad . \end{aligned} \tag{5-38}$$

Since we have $\bar{r} \times \bar{v} = \bar{h}$ we can write

$$\begin{aligned} h_z &= x\dot{y} - y\dot{x} \\ h_x &= y\dot{z} - z\dot{y} \\ h_y &= z\dot{x} - x\dot{z} \end{aligned} \tag{5-39}$$

and $$h^2 = \mu p = h_x^2 + h_y^2 + h_z^2 \qquad \mu = k^2(m_1 + m_2) \tag{5-40}$$

μ is a constant = $k^2 M_E$. From the above we can calculate p, then using

$$v^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$$

$$v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2 \tag{5-41}$$

$$r^2 = x^2 + y^2 + z^2$$

one can compute a, and from

$$p = a(1 - e^2) \tag{5-42}$$

the value of e can be obtained.

If $\bar{h}$ is projected onto the three planes, xy, yz, and zx, we obtain

$$\begin{aligned} h_z &= h \cos i \\ h_x &= \pm h \sin i \sin \Omega \\ h_y &= \mp h \sin i \cos \Omega \\ h^2 &= h_x^2 + h_y^2 + h_z^2 \quad . \end{aligned} \tag{5-43}$$

From those one obtains

$$\tan \Omega = -\frac{h_x}{h_y} \tag{5-44}$$

$$\cos i = \frac{h_z}{h} \quad . \tag{5-45}$$

The upper or lower sign is used depending upon the sign of h_z. One picks the upper value for positive h_z, corresponding to an inclination less than 90 degrees, and lower value for h_z negative, corresponding to $i > 90^\circ$.

From equations (5-31) one can obtain

$$\sin (f+\omega) = \frac{z}{r} \operatorname{cosec} i$$
$$\cos (f+\omega) = \frac{1}{r} (x \cos \Omega + y \sin \Omega) \tag{5-46}$$

from which we can solve for $\mu = f+\omega$. If $i = 0$ the equations to use are

$$\sin (f+\omega) = \frac{1}{r} (y \cos \Omega - x \sin \Omega)$$
$$\cos (f+\omega) = \frac{1}{r} (x \cos \Omega + y \sin \Omega). \tag{5-47}$$

From the relation

$$r = \frac{h^2/\mu}{1+e \cos f} \tag{5-48}$$

one can compute f and thence obtain ω.

Conversions between all kinds of coordinate systems may be found in various texts. For example, Chapter X, Page 711, of "Design Guide to Orbital Flight," by Jensen, Townsend, Kraft, and Kork, or in "Methods of Orbit Determination" by P. R. Escobal where the Appendix I (Page 393) contains 36 basic coordinate transformations.

REFERENCES

1. Brand, L., Vector Analysis, John Wiley (1957).
2. Baker, R. M. L., Jr. and M. W. Makemson, An Introduction to Astrodynamics, Academic Press (1960).
3. Ehricke, K. A., Space Flight, Vol. I, Environment and Celestial Mechanics in the series, Principles of Guided Missile Design edited by Grayson Merrill, D. Van Nostrand (1960).
4. Escobal, P. R., Methods of Orbit Determination, John Wiley (1965).
5. Jensen, J., G. Townsend, J. Kork, and D. Kraft, Design Guide to Orbital Flight, McGraw-Hill (1962).
6. Koelle, H. H. ed., Handbook of Astronautical Engineering, McGraw-Hill Book Company, Inc. (1961).
7. Roy, A. E., The Foundations of Astrodynamics, Macmillan (1965).
8. Wolaver, L. E., "A Study of Satellite Equations of Motion and Transfer Orbits," WADC Technical Report 58-378, December 1958 (AD-215854).

ELLIPTICAL FORMULAE TABLE

QUANTITY → / PAIRS ↓	a	b	r_p	r_A	e	v_c^2	v_p^2	v_A^2	$-E$	h^2
a, b	—	—	$a-\sqrt{a^2-b^2}$	$a+\sqrt{a^2-b^2}$	$\sqrt{1-\frac{b^2}{a^2}}$	$\frac{\mu}{a-\sqrt{a^2-b^2}}$	$\mu\frac{a+\sqrt{a^2-b^2}}{a^2-a\sqrt{a^2-b^2}}$	$\mu\frac{a-\sqrt{a^2-b^2}}{a^2+a\sqrt{a^2-b^2}}$	$\frac{\mu}{2a}$	$\mu\frac{b^2}{a}$
a, r_A	—	$\sqrt{2r_Aa-r_A^2}$	$2a-r_A$	—	$\frac{r_A-a}{a}$	$\frac{\mu}{2a-r_A}$	$\frac{\mu r_A}{a(2a-r_A)}$	$\mu\frac{2a-r_A}{ar_A}$	$\frac{\mu}{2a}$	$\mu\frac{2ar_A-r_A^2}{a}$
a, r_p	—	$\sqrt{2r_pa-r_p^2}$	—	$2a-r_p$	$\frac{a-r_p}{a}$	$\frac{\mu}{r_p}$	$\mu\frac{2a-r_p}{ar_p}$	$\frac{\mu}{a}\frac{r_p}{2a-r_p}$	$\frac{\mu}{2a}$	$\mu\frac{2ar_p-r_p^2}{a}$
a, e	—	$a\sqrt{1-e^2}$	$a(1-e)$	$a(1+e)$	—	$\frac{\mu}{a(1-e)}$	$\frac{\mu(1+e)}{a(1-e)}$	$\frac{\mu(1-e)}{a(1+e)}$	$\frac{\mu}{2a}$	$\mu a(1-e^2)$
b, r_p	$\frac{1}{2}\left(\frac{r_p^2+b^2}{r_p}\right)$	—	—	$\frac{b^2}{r_p}$	$\frac{b^2-r_p^2}{b^2+r_p^2}$	$\frac{\mu}{r_p}$	$\frac{2\mu b^2}{r_p(r_p^2+b^2)}$	$\frac{r_p^2}{b^2}\frac{2\mu r_p}{r_p^2+b^2}$	$\mu\frac{r_p}{r_p^2+b^2}$	$\mu\frac{2r_pb^2}{b^2+r_p^2}$
b, r_A	$\frac{1}{2}\left(\frac{r_A^2+b^2}{r_A}\right)$	—	$\frac{b^2}{r_A}$	—	$\frac{r_A^2-b^2}{r_A^2+b^2}$	$\frac{\mu r_A}{b^2}$	$\frac{2\mu r_A^2}{b^2(r_A^2+b^2)}$	$\frac{2\mu b^2}{r_A(b^2+r_A^2)}$	$\frac{\mu r_A}{r_A^2+b^2}$	$\mu\frac{2r_Ab^2}{r_A^2+b^2}$
b, e	$\frac{b}{\sqrt{1-e^2}}$	—	$b\sqrt{\frac{1-e}{1+e}}$	$b\sqrt{\frac{1+e}{1-e}}$	—	$\frac{\mu\sqrt{1+e}}{b\sqrt{1-e}}$	$\frac{\mu\sqrt{1-e^2}}{b}\left(\frac{1+e}{1-e}\right)$	$\frac{\mu\sqrt{1-e^2}}{b}\left(\frac{1-e}{1+e}\right)$	$\mu\frac{\sqrt{1-e^2}}{2b}$	$\mu\frac{b(1-e^2)}{\sqrt{1-e^2}}$
r_p, r_A	$\frac{1}{2}(r_A+r_p)$	$\sqrt{r_Ar_p}$	—	—	$\frac{r_A-r_p}{r_A+r_p}$	$\frac{\mu}{r_p}$	$\frac{2\mu r_A}{r_p(r_A+r_p)}$	$\frac{2\mu r_p}{r_A(r_p+r_A)}$	$\frac{\mu}{r_A+r_p}$	$\mu\frac{2r_Ar_p}{r_A+r_p}$
r_A, e	$\frac{r_A}{1+e}$	$r_A\sqrt{\frac{1-e}{1+e}}$	$\frac{r_A(1-e)}{1+e}$	—	—	$\frac{\mu(1+e)}{r_A(1-e)}$	$\frac{\mu(1+e)^2}{r_A(1-e)}$	$\frac{\mu}{r_A}(1-e)$	$\frac{\mu(1+e)}{2r_A}$	$\mu r_A(1-e)$
r_p, e	$\frac{r_p}{1-e}$	$r_p\sqrt{\frac{1+e}{1-e}}$	—	$\frac{r_p(1+e)}{1-e}$	—	$\frac{\mu}{r_p}$	$\frac{\mu}{r_p}(1+e)$	$\frac{\mu}{r_p}\frac{(1-e)^2}{1+e}$	$\frac{\mu(1-e)}{2r_p}$	$\mu r_p(1+2)$
r_p, v_p	$\frac{\mu r_p}{2\mu-r_pv_p^2}$	$r_pv_p\sqrt{\frac{r_p}{2\mu-r_pv_p^2}}$	—	$\frac{r_p^2v_p^2}{2\mu-r_pv_p^2}$	$\frac{r_pv_p^2}{\mu}-1$	$\frac{\mu}{r_p}$	—	$\left(\frac{2\mu-r_pv_p^2}{r_pv_p}\right)^2$	$\frac{2\mu-r_pv_p^2}{2r_p}$	$r_p^2v_p^2$
r_A, v_A	$\frac{\mu r_A}{2\mu-r_Av_A^2}$	$\sqrt{\frac{r_A^3v_A^2}{2\mu-r_Av_A^2}}$	$\frac{r_A^2v_A^2}{2\mu-r_Av_A^2}$	—	$1-\frac{r_Av_A^2}{\mu}$	$2\left(\frac{\mu}{r_Av_A}\right)^2-\frac{\mu}{r_A}$	$\left(\frac{2\mu-r_Av_A^2}{r_Av_A}\right)^2$	—	$\frac{\mu}{r_A}-\frac{1}{2}v_A^2$	$r_A^2v_A^2$
$h, +E$	$-\frac{\mu}{2E}$	$\frac{h}{\sqrt{-2E}}$	$-\frac{\mu}{2E}\left[1-\sqrt{1+\frac{2Eh^2}{\mu^2}}\right]$	$-\frac{\mu}{2E}\left[1+\sqrt{1+\frac{2Eh^2}{\mu^2}}\right]$	$\sqrt{1+\frac{2Eh^2}{\mu^2}}$	$\frac{-2E}{1-\sqrt{1+\frac{2Eh^2}{\mu^2}}}$	$-2E\,\frac{\mu+\sqrt{\mu^2+2Eh^2}}{\mu-\sqrt{\mu^2+2Eh^2}}$	$-2E\,\frac{\mu-\sqrt{\mu^2+2Eh^2}}{\mu+\sqrt{\mu^2+2Eh^2}}$	—	—

$p = r_p(1+e) = a(1-e^2) = \frac{b^2}{a}$

$h = r_p v_p = r_A v_A$ = ANGULAR MOMENTUM

$h^2 = \mu p$

$E = \frac{1}{2}v_p^2 - \frac{\mu}{r_p} = \frac{1}{2}v_A^2 - \frac{\mu}{r_A} = -\frac{\mu}{2a}$ = ENERGY

$v^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right)$ = (VIS-VIVA)

$v_A^2 = \frac{v_p^2(1-e)^2}{(1+e)^2}$

$v_c^2 = \frac{\mu}{r_p}$

$\mu = k^2(m_1+m_2)$

k^2 = GAUSS' CONSTANT

(PAGE 55)

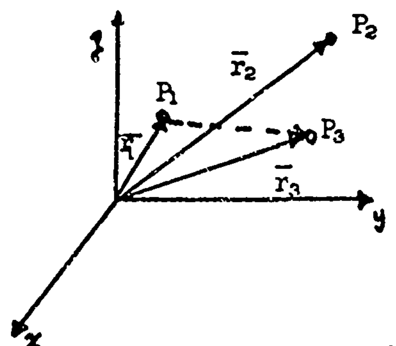
LEW/ARL/1970

FILL IN THE BLANKS -

QUANTITY → PAIRS ↓	a	b	r_p	r_A	e	v_p^2	v_A^2	$-E$	h^2	p	v_c^2
a, p	—	$\sqrt{ap}$								—	
r_p, p			—							—	
r_A, p				—						—	
e, p					—					—	
v_c, p										—	—
e, r_p	$\frac{r_p}{1-e}$		—		—	$\frac{\mu}{r_p}(1+e)$			$\mu r_p(1+e)$		$\frac{\mu}{r_p}$
e, v_p					—	—					
e, v_A					—		—				
e, r_A	$\frac{r_A}{1+e}$			—	—						
r_p, c	$\frac{r_p}{2-c^2}$		—		c^2-1						

$c^2 = \frac{v_p^2}{v_c^2}$ $\quad v_c = \sqrt{\frac{\mu}{r_p}}$ $\quad p = r_A(1-e) = r_p(1+e)$ $\quad r_p = a(1-e)$ $\quad r_A = a(1+e)$

6. THREE AND N-BODY PROBLEMS



$$\overline{P_1P_3} = \bar{r}_3 - \bar{r}_1$$

FIGURE 6-1

For three bodies the equations of motion are

$$m_1 \frac{d^2\bar{r}_1}{dt^2} = k^2 \frac{m_1 m_2 (\bar{r}_2 - \bar{r}_1)}{r_{12}^3} + k^2 \frac{m_1 m_3 (\bar{r}_3 - \bar{r}_1)}{r_{13}^3}$$

$$m_2 \frac{d^2\bar{r}_2}{dt^2} = k^2 \frac{m_2 m_1 (\bar{r}_1 - \bar{r}_2)}{r_{21}^3} + k^2 \frac{m_2 m_3 (\bar{r}_3 - \bar{r}_2)}{r_{23}^3} \qquad (6\text{-}1)$$

$$m_3 \frac{d^2\bar{r}_3}{dt^2} = k^2 \frac{m_3 m_1 (\bar{r}_1 - \bar{r}_3)}{r_{31}^3} + k^2 \frac{m_3 m_1 (\bar{r}_2 - \bar{r}_3)}{r_{32}^3}$$

In the case of n bodies, we have equations of the form,

$$m_i \frac{d^2 x_i}{dt^2} = - k^2 m_i \sum_{j=0}^{n-1} m_j \frac{(x_i - x_j)}{r_{ij}^3}$$

$$m_i \frac{d^2 y_i}{dt^2} = - k^2 m_i \sum_{j=0}^{n-1} m_j \frac{(y_i - y_j)}{r_{ij}^3} \qquad (6\text{-}2)$$

$$m_i \frac{d^2 z_i}{dt^2} = - k^2 m_i \sum_{j=0}^{n-1} m_j \frac{(z_i - z_j)}{r_{ij}^3}$$

$i = 1, 2, \dots n \quad j \neq i$

When we consider the three body problem in rectangular coordinates, there are 9 equations of second order for a total of a system of equations of 18th order. These have been effectively unsolved in general. In order to solve these problems we look for "integrals" or "first integrals" or "constants of the motion." These are basically conservation theorems.

In general for a system

$$\frac{dx_i}{dt} = F_i \ (x_1, x_2, \ldots x_n, t) \quad i = 1, 2, \ldots n$$

we seek an "integral," a function of the form,

$$\varphi \ (x_1, x_2, \ldots x_n, t)$$

such that

$$\frac{d\varphi}{dt} = 0$$

on each solution of F_i.

When n = 2 we have solutions as paths in this space. $d\varphi/dt = 0$ means that φ = constant on each curve.

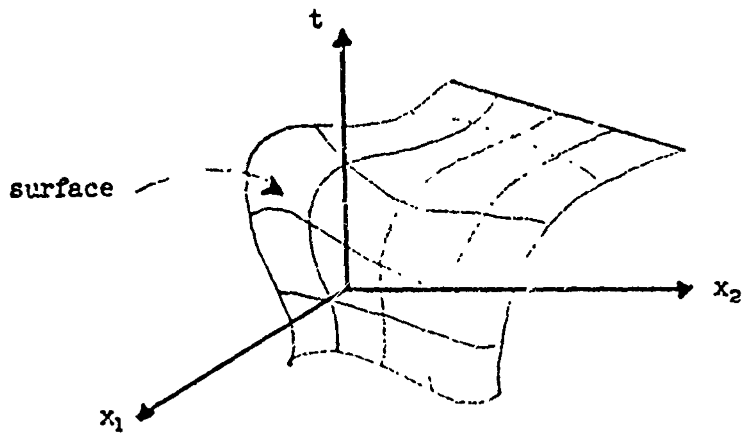


FIGURE 6-2

Take $\phi = 1$, say, a surface in the 3-D space. This surface is built up of complete solution curves. Each level surface of ϕ, ($\phi = 1, 2, \ldots$ for many surfaces) is made up of complete solution curves. All of the level surfaces form a stratification. Since ϕ is a surface and therefore 2-D, we have reduced the problem by one dimension.

We use the relation $\phi\ (x_1, x_2, \ldots x_n, t) = C$ to eliminate one variable and thereby reduce the order by one. These ϕ formulae are conservative laws. In general, if we have k such integrals, $\phi_1, \phi_2, \ldots \phi_k$, then the system order can be lowered by the order k.

$$\begin{array}{ccc} \phi_1\ (x_1, x_2, \ldots x_n, t) & = & C_1 \\ \vdots & & \vdots \\ \phi_k\ (x_1, x_2, \ldots x_n, t) & = & C_k \end{array}$$

However, to solve these we must have at least one Jacobian (page 115) of order $k \neq 0$. That is, $\partial(\phi_1, \phi_2, \ldots \phi_k)/\partial(x_1, x_2, \ldots x_k) \neq 0$ at least in a region. When this is satisfied we can solve for $x_1, x_2, \ldots x_k$ in terms of the other variables to lower the system order by k.

We have a general theorem in celestial mechanics problems which says these integrals <u>must exist</u>. The third dimension is time, which increases monotonically, the solution can't turn around in time.

For the three body problem only ten of the required 18 integrals are known. Let us consider the known integrals.

ENERGY

$$m_1 \frac{d^2\bar{r}_1}{dt^2} = -\text{grad}_1 U; \; m_2 \frac{d^2 r_2}{dt^2} = -\text{grad}_2 U$$

$$U = -k^2 \left[\frac{m_1 m_2}{r_{12}} + \frac{m_2 m_3}{r_{23}} + \frac{m_1 m_3}{r_{13}} \right]$$

(1) $$\sum_{i=1}^{3} \frac{1}{2} m_i \left| \frac{d\bar{r}_i}{dt} \right|^2 + U = C_1 = \text{constant}$$

For n bodies this takes the form

$$\frac{1}{2} \sum_{i=1}^{n} m_i \left\{ \left(\frac{dx_i}{dt}\right)^2 + \left(\frac{dy_i}{dt}\right)^2 + \left(\frac{dz_i}{dt}\right)^2 \right\} + U = C_1 = \text{constant}$$

$$U = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{k^2 m_i m_j}{r_{ij}} \qquad i \neq j$$

For the case of internal forces only, the conservation of linear momentum holds.

(3) $\sum_{i=1}^{n} m_i \frac{d\bar{r}_i}{dt} = \bar{k}$ with three components of $\bar{k}$ we have three equations, therefore three "integrals."

The center of mass moves in a straight line.

(3) $$\sum_{i=1}^{n} m_i \bar{r}_i = \bar{k}t + \bar{l}$$

The conservation of angular momentum gives three more

(3) $$\sum_{i=1}^{n} m_i \bar{r}_i \times \frac{d\bar{r}_i}{dt} = \bar{h}$$

These may be written as

$$
\begin{aligned}
&\sum_{i=1}^{n} m_i \left[x_i \frac{dy_i}{dt} - y_i \frac{dx_i}{dt}\right] = C_1 \\
&\sum_{i=1}^{n} m_i \left[y_i \frac{dz_i}{dt} - z_i \frac{dy_i}{dt}\right] = C_2 \qquad\qquad (6\text{-}3) \\
&\sum_{i=1}^{n} m_i \left[z_i \frac{dx_i}{dt} - x_i \frac{dz_i}{dt}\right] = C_3 .
\end{aligned}
$$

LaPlace showed that one could direct the axes so that two of the constants in these equations, say C_1 and C_2 would be zero while the third becomes $\sqrt{C_1^2 + C_2^2 + C_3^2}$. This is the plane of maximum sum of the products of the masses and the rates of the projection of areas. It is called the invariable plane by LaPlace. The invariable plane of the solar system is inclined to the ecliptic by about 2 degrees with Ω = 286 degrees (see Chapter 32).

Thus we need 6n integrals for the n body problem but can find only ten. Can we find any more? Bruns (Acta Mathematics Vol XI) showed that with rectangular coordinates no new algebraic integrals are possible. Poincare' showed there were no new uniform transcendental integrals when the coordinates are the elements of the orbits. So we can't solve even the three body problem in general; however, for special cases there are some singular points.

Before discussing these singular points let us develop the geocentric form of the n body problem. If we place the origin of a fixed inertial coordinate system at the center of mass of the Solar System, the equations of motion in component form become

$$\frac{d^2\xi}{dt^2} = k^2 \sum_i m_i \frac{(\xi_i - \xi)}{r_i^{\,3}} \tag{6-4}$$

$$\frac{d^2\eta}{dt^2} = k^2 \sum_i m_i \frac{(\eta_i - \eta)}{r_i^{\,3}}$$

$$\frac{d^2\zeta}{dt^2} = k^2 \sum_i m_i \frac{(\zeta_i - \zeta)}{r_i^{\,3}}$$

$$r_i^{\,2} = (\xi_i - \xi)^2 + (\eta_i - \eta)^2 + (\zeta_i - \zeta)^2 \tag{6-5}$$

where ξ_i, η_i, ζ_i are the coordinates of the planets of mass m_i and ξ, η, ζ are the coordinates of the particle whose motion is being investigated. It is desirable to move the coordinate system to the center of mass of the Earth. To do this, let

$$\begin{aligned} x &= \xi - \xi_o \\ y &= \eta - \eta_o \\ z &= \zeta - \zeta_o \end{aligned} \tag{6-6}$$

where ξ_o, η_o, ζ_o are the coordinates of the Earth with respect to the center of mass of the solar system, for which

$$\frac{d^2\xi_o}{dt^2} = k^2 \sum_{i=1}^{n-1} m_i \frac{(\xi_i - \xi_o)}{r_{io}^{\,3}} \quad , \quad \text{etc.} \tag{6-7}$$

r_{io} = distance from center of Earth to planet with mass m_i.

Similar equations exist for the other two components η_o and ζ_o. We now form

$$\frac{d^2\xi}{dt^2} - \frac{d^2\xi_0}{dt^2} = k^2 m_0 \frac{\xi_0 - \xi}{r_0^3} + k^2 \sum_{i=1}^{n-1} m_i \left[\frac{\xi_i - \xi}{r_i^3} - \frac{(\xi_i - \xi_0)}{r_{i0}^3} \right] \qquad (6\text{-}8)$$

from which we obtain

$$\frac{d^2x}{dt^2} = -\frac{k^2 m_0 x}{r^3} + k^2 \sum_{i=1}^{n-1} m_i \left[\frac{x_i - x}{r_i^3} - \frac{x_i}{r_{i0}^3} \right] \qquad (6\text{-}9)$$

and the equations for the other coordinates are derived similarly.

$$r^2 = x^2 + y^2 + z^2 \qquad (6\text{-}10)$$

$$r_i^2 = (x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2 = (\text{particle to planet distance})^2$$

$$r_{i0}^2 = x_i^2 + y_i^2 + z_i^2 = (\text{Earth to planet distance})^2$$

The values of x_i, y_i, z_i are known tabulated functions of time obtained from the American Ephemeris and Nautical Almanac. Note the equations of motion for the effect of the particle on the planets have been neglected. We will have occasion to return to equation (6-9) many times. Since x, y, and z are measured from the center of the Earth, the system is called the geocentric coordinate system.

The n-body problem is pretty hopeless to solve. Given three finite bodies a general solution cannot be found; however, there are particular solutions to the three body problem. If each of the bodies is placed at the vertices of an equilateral triangle, they will maintain that configuration, although the triangle itself may vary in size. In addition there is a straight line configuration. These solutions were found by Lagrange in 1772. The figure below gives examples of this periodic motion.

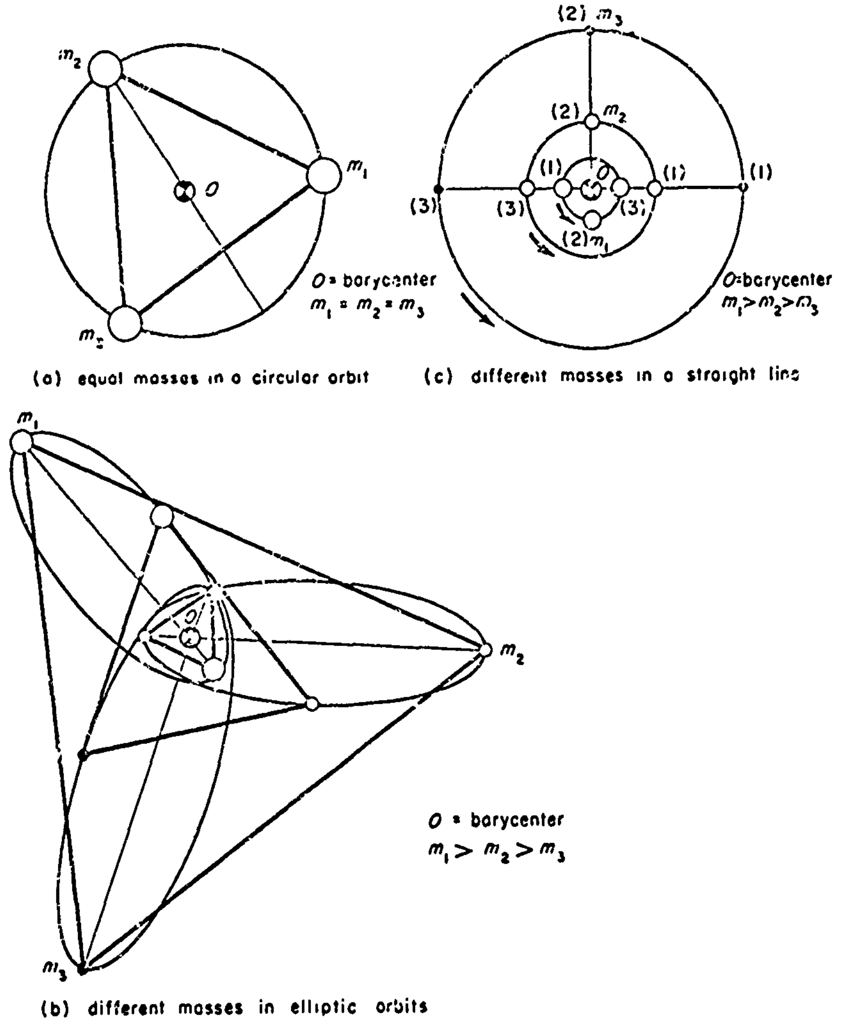
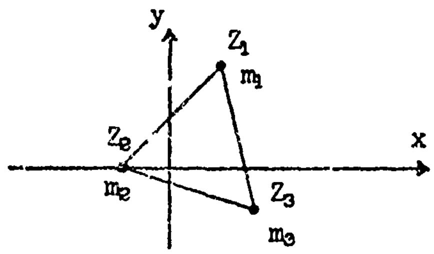


FIG. 6-3 Examples for Lagrange's Periodic Solutions of the Three-Body Problem.

FIGURE 6-3

Now let's consider the analysis. In the general case the solution to the three body problem is planar. This fact is shown by Wintner in his book, "The Analytical Foundations of Celestial Mechanics." Let's assume planarity and consider the possible solutions in a fixed plane with the origin at the center of mass. To aid the solution we consider complex numbers Z.



$$m_1 Z_1 + m_2 Z_2 + m_3 Z_3 = 0 \tag{6-11}$$

since the origin is the center of mass.

$$Z_3 - Z_1 = \omega(Z_2 - Z_1) \tag{6-12}$$

where $\omega = e^{i\,\pi/3}$, is the equation for the condition that we have an equilateral triangle. Equations (6-11) and (6-12) give

$$Z_2 = Z_1 \left[1 - \frac{m_1 + m_2 + m_3}{m_2 + \omega m_3}\right] = Z_1 (1 + a)$$

$$Z_3 = Z_1 \left[1 - \omega \frac{m_1 + m_2 + m_3}{m_2 + m_3 \omega}\right] = Z_1 (1 + a\omega) \tag{6-13}$$

where $a = - \dfrac{m_1 + m_2 + m_3}{m_2 + \omega m_3}$ = complex number.

Now by Newton's inverse square law we have

$$m_1 \frac{d^2 Z_1}{dt^2} = k^2 \frac{m_1 m_2 (Z_2 - Z_1)}{|Z_2 - Z_1|^3} + \frac{k^2 m_1 m_3 (Z_3 - Z_1)}{|Z_3 - Z_1|^3}. \tag{6-14}$$

For an equilateral triangle,

$$|Z_2 - Z_1| = |Z_3 - Z_1| = |Z_3 - Z_2| = \rho = |a| \; |Z_1| \tag{6-15}$$

Where a is a complex number and $Z_2 - Z_1 = a Z_1$; $Z_3 - Z_1 = a\omega Z_1$.

Hence:

$$m_1 \frac{d^2 Z_1}{dt^2} = k^2 m_2 \frac{a Z_1 m_1}{|a|^3 \, |Z_1|^3} + \frac{k^2 m_3 a\omega Z_1 m_1}{|a|^3 \, |Z_1|^3}$$

$$\frac{d^2 Z_1}{dt^2} = \frac{Z_1 k^2}{|a|^3 \, |Z_1|^3} \left[\frac{-m_2 (m_1 + m_2 + m_3)}{m_2 + m_3 \omega} - \frac{m_3 \omega (m_1 + m_2 + m_3)}{m_2 + m_3 \omega} \right]$$

$$\frac{d^2 Z_1}{dt^2} = -\frac{k^2 Z_1 (m_1 + m_2 + m_3)}{|a|^3 \, |Z_1|^3} \tag{6-16}$$

which is our central force inverse square law. Therefore, if there is a triangle solution, then each particle will move in a conic section in similar orbits. Hence for any solution of this type, Z_1, Z_2 and Z_3 all satisfy,

$$\frac{d^2 Z}{dt^2} = -\frac{k^2 m \, Z}{|a|^3 \, |Z|^3} \, . \tag{6-17}$$

The three orbits are similar because of the relations:

$$Z_2 = (1 + a)Z_1 \quad \text{and} \quad Z_3 = (1 + a\omega)Z_1 \, . \tag{6-18}$$

Now let us examine the converse property. Let Z_1 move in any solution of this central inverse square force field and place Z_2 and Z_3 at $(1 + a)Z_1$ and $(1 + a\omega)Z_1$ respectively, then we get a solution of the three body problem. The equation for Z_1 is, of course, the same as before. For Z_2 we must verify that,

$$m_2 \frac{d^2 Z_2}{dt^2} = \frac{k^2 m_1 (Z_1 - Z_2) m_2}{|Z_1 - Z_2|^3} + \frac{k^2 m_3 (Z_3 - Z_2) m_2}{|Z_3 - Z_2|^3}$$

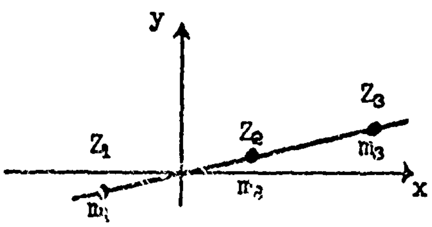
$$\frac{d^2 Z_2}{dt^2} = (1 + a)\frac{d^2 Z_1}{dt^2} = \frac{k^2 m_1 (-aZ_1)}{|a|^3 \, |Z_1|^3} + \frac{k^2 m_3 (a\omega - a) Z_1}{|a|^3 \, |Z_1|^3}$$

$$(1 + a)\frac{d^2 Z_1}{dt^2} = (1 + a)\,\frac{-k^2 m \, Z_1}{|a|^3 \, |Z_1|^3} \tag{6-19}$$

Hence such motion for Z_2 is possible. Similarly for Z_3. It is interesting to note that this kind of motion is possible for any central force field with power laws other than inverse square law since the factor $|a|^3 |Z_1|^3$ appears throughout the equations.

We have thus established the existence of the sextile points as a solution to the three body equations even when all three bodies have commeasurable masses and all move in any conic orbit.

Now let's look at another class of solutions to the same problem, the straight line solutions. We desire solutions such that the three finite bodies all lie on a straight line for all time. (The line is, of course, rotating however). Again the solution can be shown to be planar, and for motion in a plane we take the center of mass as the origin of the axis system in the Z plane just as before.



Since the origin is the center of mass, we have

$$m_1 Z_1 + m_2 Z_2 + m_3 Z_3 = 0 \tag{6-20}$$

Assume a straight line motion. Then using p and q as constants, (real numbers)

$0 \leq p < q$

this straight line relationship is represented mathematically as

$$Z_2 = -pZ_1 \qquad Z_3 = -qZ_1 \tag{6-21}$$

$$\frac{Z_2 - Z_1}{Z_3 - Z_2} = \frac{-pZ_1 - Z_1}{-qZ_1 + pZ_1} = \frac{1+p}{q-p} = b = \text{real number.}$$

Having the ratio as a constant real number means we have straight line motion. Assume this fixed ratio, b. Knowing that the straight line solutions are such that the ratio of the distances are constant, we assume this and try to solve in reverse to how that

$$b = \frac{Z_2 - Z_1}{Z_3 - Z_2} \; . \tag{6-22}$$

Now,

$$\frac{d^2 Z_1}{dt^2} = \frac{k^2 m_2 (Z_2 - Z_1)}{|Z_2 - Z_1|^3} + \frac{k^2 m_3 (Z_3 - Z_1)}{|Z_3 - Z_1|^3}$$

$$\frac{d_2 Z_1}{dt^2} = \frac{-k^2 m_2 (1+p) Z_1}{(1+p)^3 \, |Z_1|^3} - \frac{k^2 m_3 (1+q) Z_1}{(1+q)^3 \, |Z_1|^3} = -A \frac{Z_1}{|Z_1|^3} \tag{6-23}$$

where

$$A = \frac{k^2 m_2}{(1+p)^2} + \frac{k^2 m_2}{(1+q)^2} \; . \tag{6-24}$$

Each body moves in a central force field, i.e., motion is a conic section. Then we must determine if Z_2 and Z_3 will satisfy the differential equation. Now because the center of mass is at the origin we have

$$m_1 Z_1 + m_2 Z_2 + m_3 Z_3 = 0$$

$$m_1 Z_1 - m_2 p Z_1 - m_3 q Z_1 = 0$$

From which we find

$$m_1 = m_2 p + m_3 q \tag{6-25}$$

Now let Z_1 move in a solution of its equation,

$$\ddot{Z}_1 = -A \frac{Z}{|Z_1|^3} \tag{6-26}$$

and put Z_2 at $-pZ_1$ and Z_3 at $-qZ_1$ and then require that the differential equation be satisfied. The equation for Z_1 is, of course, automatically satisfied. Now examine the motion of Z_2.

$$\frac{d^2Z_2}{dt^2} = k^2 \frac{m_1(Z_1 - Z_2)}{|Z_1 - Z_2|^3} + \frac{k^2 m_3(Z_3 - Z_2)}{|Z_3 - Z_2|^3}$$

$$-p \frac{d^2Z_1}{dt^2} = \frac{k^2 m_1 Z_1(1+p)}{(1+p)^3 |Z_1|^3} + \frac{k^2 m_3 Z_1(p-q)}{(|p-q|^3)|Z_1|^3} \tag{6-27}$$

Knowing that

$$\frac{d^2Z_1}{dt^2} = -A \frac{Z_1}{|Z_1|^3} \tag{6-28}$$

and

$$m_1 = m_2 p + m_3 q \quad , \quad q > p, \tag{6-29}$$

we have, after using Equation (6-24) for A,

$$Ap \frac{Z_1}{|Z_1|^3} = pk^2 \left[\frac{m_2}{(1+p)^2} + \frac{m_3}{(1+q)^2}\right] \frac{Z_1}{|Z_1|^3} = \frac{k^2 m_1 Z_1}{(1+p)^2 |Z_1|^3} - \frac{k^2 m_3 Z_1}{(q-p)^2 |Z_1|^3}$$

$$p \left[\frac{m_2}{(1+p)^2} + \frac{m_3}{(1+q)^2}\right] = \frac{m_2 p + m_3 q}{(1+p)^2} - \frac{m_3}{(q-p)^2} = \frac{pm_2}{(1+p)^2} + \frac{qm_3}{(1+p)^2} - \frac{m_3}{(q-p)^2}$$

$$\frac{\cancel{m}_3 p}{(1+q)^2} = \frac{q\cancel{m}_3}{(1+p)^2} - \frac{\cancel{m}_3}{(q-p)^2} \tag{6-30}$$

We must have:

$$\frac{p}{(1+q)^2} = \frac{q}{(1+p)^2} - \frac{1}{(q-p)^2} \tag{6-31}$$

The equation for Z_2 leads to the same conditions. Thus for arbitrary masses we can pick p and q subject to $0 \le p < q$ and have a unique solution of the equations as a straight line solution. Hence, given that

$$Z_2 = -pZ_1 \qquad Z_3 = -qZ_1 \qquad 0 \le p < q \tag{-32}$$

and given three finite masses m_1, m_2 and m_3 we can choose p and q to satisfy

$$m_1 - pm_2 - qm_3 = 0 \tag{6-33}$$

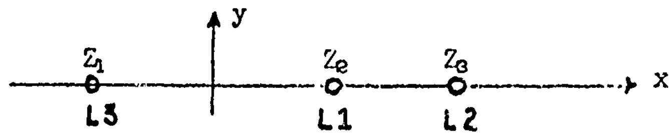
and

$$\frac{q}{(1+p)^2} - \frac{1}{(q-p)^2} - \frac{p}{(1+q)^2} = 0 \tag{6-34}$$

This gives rise to a quintic equation for p and q. Now one can ask if it is always possible to satisfy the two equations for any mass ratio. It can be shown that it is necessary that the largest of the three masses be on the negative side of the center of mass from the larger of the remaining two masses. This is really no restriction, just a note on the orientation of the axis system.

Note that the solution holds for three finite masses and for any conic section motion they assume. Now let us further restrict the problem by assuming one of the masses to be negligible compared to the other two masses such that it does not disturb their orbits. We still assume elliptical motion, however. Let the distance between the two massive bodies be unity.

For convenience let's assume the configuration below:



For the L_1 point we have $m_2 = 0$ and let $m_1 > m_3$ $\quad M = m_3/m_1$.
Then the Equations (6-33) and (6-34) on the above become:

$$q = \frac{m_1}{m_3} = \frac{1}{M} \tag{6-35}$$

$$\frac{q}{(1+p)^2} - \frac{1}{(q-p)^2} - \frac{p}{(1+q)^2} = 0 \ . \tag{6-36}$$

Substituting from (6-35) into (6-36) gives

$$\frac{1}{M(1+p)^2} - \frac{1}{\left(\frac{1}{M} - p\right)^2} - \frac{p}{\left(1 + \frac{1}{M}\right)^2} = 0$$

$$\frac{1}{M(1+p)^2} - \frac{M^2}{(1-Mp)^2} - \frac{pM^2}{(M+1)^2} = 0 \tag{6-37}$$

which reduces to the quintic in p. Now,

$$Z_1 = -\mu \qquad Z_2 = r_1 \qquad Z_3 = 1-\mu \qquad \mu = \frac{M}{1+M}$$

$$p = -\frac{Z_2}{Z_1} = \frac{r_1}{\mu} \qquad M = \frac{\mu}{1-\mu} \ . \tag{6-38}$$

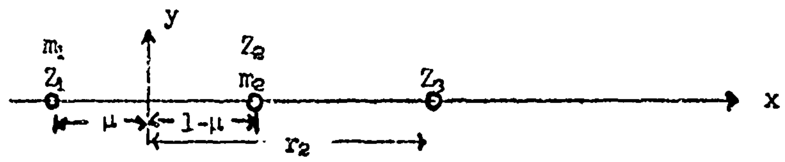
In terms of r_1 the distance of L_1 from the origin at the center of mass, Equation (6-37) becomes,

$$\frac{1-\mu}{(\mu+r_1)^2} - \frac{\mu}{(1-\mu-r_1)^2} - r_1 = 0 \tag{6-39}$$

$$r_1 = \frac{1-\mu}{(\mu+r_1)^2} - \frac{\mu}{(1-\mu-r_1)^2} \tag{6-40}$$

We can use this to iterate to a correct value for r_1. However, the problem has been quite accurately solved by K. D. Abhyankar. His results are given in Table I.

Similarly for L_2 we have that



$$m_3 = 0 \qquad p = \frac{m_1}{m_2} = \frac{1}{M} \qquad M = \frac{m_2}{m_1}$$

and Equation (6-36) becomes

$$\frac{q}{\left(1+\frac{1}{M}\right)^2} - \frac{1}{\left(q-\frac{1}{M}\right)^2} - \frac{1}{M(1+q)^2} = 0$$

$$\frac{M^2 q}{(1+M)^2} - \frac{M^2}{(Mq-1)^2} - \frac{1}{M(1+q)^2} = 0 \tag{6-41}$$

which gives a quintic in q.

Now

$$Z_1 = -\mu \qquad Z_2 = 1-\mu \qquad Z_3 = r_2$$

$$q = -\frac{Z_3}{Z_1} = \frac{r_2}{\mu}.$$

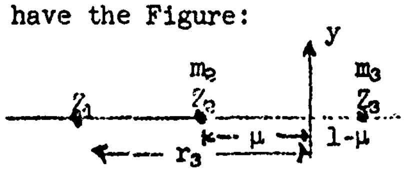
In terms of r_2, the distance of L_2 from the origin (which is at the center of mass), we have, after substitution for q from Equation (6-41)

$$r_2 - \frac{\mu}{(r_2-1+\mu)^2} - \frac{1-\mu}{(r_2+\mu)^2} = 0 \tag{6-42}$$

which is of the form

$$r_2 = \frac{1-\mu}{(r_2+\mu)^2} + \frac{\mu}{(r_2-1+\mu)^2} \quad . \tag{6-43}$$

This can be used to iterate for the value of r_2. Now for the case of L_3 we have the Figure:



$$m_1 = 0 \qquad m_2 > m_3$$

$$M = \frac{m_3}{m_2} = \frac{\mu}{1-\mu} \qquad \mu = \frac{M}{1+M}$$

$$p_2 = -\frac{Z_2}{Z_1} \qquad q = -\frac{Z_3}{Z_1}$$

$$Z_2 = -\mu \qquad Z_3 = 1-\mu \qquad Z_1 = -r_3$$

$$p = -\frac{-\mu}{-r_3} = -\frac{\mu}{r_3} \qquad\qquad q = +\frac{1-\mu}{+r_3}$$

Equation (6-33) is automatically satisfied. Equation (6-34) becomes

$$\frac{q}{(1+p)^2} - \frac{1}{(q-p)^2} - \frac{p}{(1+q)^2} = 0 \tag{6-44}$$

substituting

$$p = \frac{-\mu}{r_3} \qquad q = \frac{1-\mu}{r_3}$$

gives

$$\frac{1-\mu}{(r_3-\mu)^2} - r_3 + \frac{\mu}{(r_3+1-\mu)^2} = 0 \tag{6-45}$$

which is of the form

$$r_3 = \frac{\mu}{(r_3+1-\mu)^2} + \frac{1-\mu}{(r_3-\mu)^2} \tag{6-46}$$

These values of r_1, r_2, r_3 as found by Abhyankar are given in Table I on the next page.

REFERENCES

Ehricke, K. A., Space Flight Vol I. Environment and Celestial Mechanics, D. Van Nostrand (1960).

American Ephemeris and Nautical Almanac. U.S. Government Printing Office, Washington D. C.

Danby, J. M. A., Fundamentals of Celestial Mechanics, Macmillan (1962).

Poincaré, H., Les Méthodes Nouvelles de la Mécanique Céleste, (3 vols.) Gauthier-Villars, Paris (1895), reissued by Dover.

Roy, A. E., The Foundations of Astrodynamics, Macmillan Company (1965).

Kaplan, W., An Introduction to Celestial Mechanics, University of Michigan, Institute of Science and Technology Report, April 1961.

Wintner, A., Analytical Foundations of Celestial Mechanics, Princeton Press (1941).

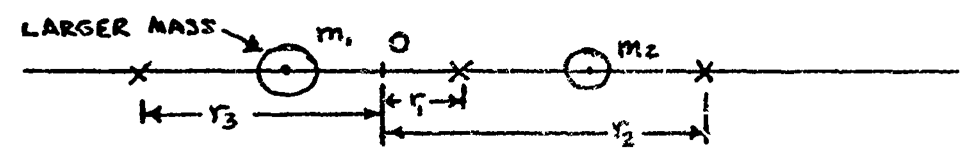


TABLE I STRAIGHT LINE SOLUTION INITIAL DISTANCE

u	M	r_1	r_2	r_3
0.5000 0000	1.0	0.0000 00000	1.1984 06144	1.1984 06144
0.4444 4444	4/5	0.0785 05894	1.2170 25950	1.1785 83395
0.4000 0000	2/3	0.1416 17526	1.2308 13769	1.1620 45267
0.3750 0000	3/5	0.1773 42630	1.2380 38366	1.1525 24009
0.3000 0000	3/7	0.2861 29783	1.2567 34696	1.1232 05596
0.2857 1429	2/5	0.3072 33226	1.2596 66516	1.1175 14574
0.2500 0000	1/3	0.3607 43428	1.2658 58102	1.1031 66848
0.2307 6923	3/10	0.3900 97486	1.2684 02205	1.0953 78413
0.2000 0000	1/4	0.4380 75959	1.2710 48690	1.0828 39465
0.1500 0000	3/17	0.5197 40359	1.2703 34073	1.0622 98631
0.1000 0000	1/9	0.6090 35110	1.2596 99833	1.0416 08909
0.0500 0000	1/19	0.7152 25350	1.2280 93667	1.0208 26334
0.0200 0000	1/49	0.8034 65629	1.1800 77904	1.0083 32894
0.0100 0000	1/99	0.8480 78713	1.1467 65042	1.0041 66612
0.0030 0000		0.9003 73655	1.1002 85368	1.0012 49999
0.0010 0000		0.9312 86976	1.0699 16097	1.0004 16667
0.0003 0000		0.9540 10057	1.0468 25818	1.0001 25000
0.0001 0000		0.9680 65213	1.0324 25189	1.0000 41667
0.0000 3000		0.9785 81345	1.0216 68072	1.0000 12501
0.0000 1000		0.9851 26720	1.0150 02015	1.0000 04168
0.0000 0300		0.9900 30459	1.0100 30190	1.0000 01251
0.0000 0100		0.9930 81618	1.0069 48468	1.0000 00417
0.0000 0030		0.9953 65315	1.0046 48045	1.0000 00125
0.0000 0010		0.9967 85658	1.0032 21637	1.0000 00042
0		1.0000 00000	1.0000 00000	1.0000 00000

$$M = \frac{m_2}{m_1} < 1.0$$

$$\mu = \frac{M}{1+M} = \frac{m_2}{m_1+m_2}$$

7. RESTRICTED THREE BODY PROBLEM

The three body problem may be further reduced by considering the case where one of the three masses is so small that it does not disturb the orbit of the other two. These two can then be solved separately - a two body problem - and then their effect on the small body investigated. We can find one additional integral by this means. Thus the third body moves subject to P1 and P2 but not affecting them. Since the total force on the third body is not now directed toward a fixed center, the $r^2\dot{f}$ term is <u>not</u> conserved and the force-field varies with time. A consequence of this is that energy is <u>not</u> conserved in the rotating system of coordinates. The absence of these integrals makes the restricted three body problem more difficult than the two body case. However, if we assume a circular orbit for the two large masses then another integral may be found.

To be precise, the restricted three body problem refers to one very small body with the two large bodies moving in a <u>circular</u> orbit about their common center of mass.

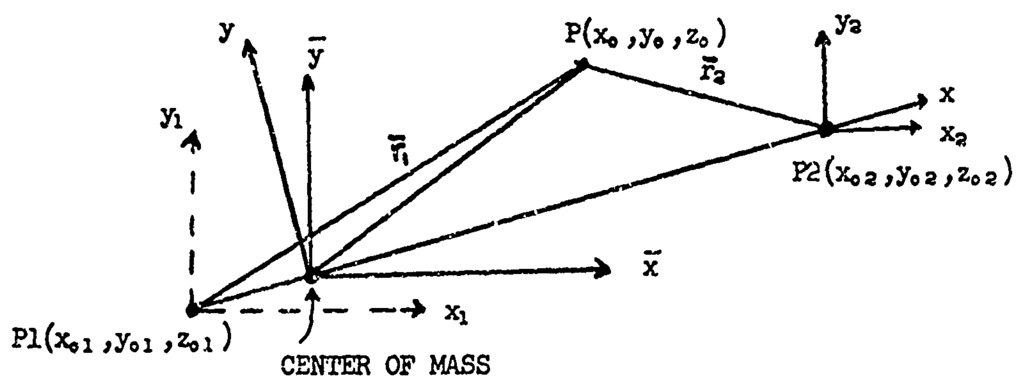


FIGURE 7-1

Assume P1 and P2 move in a circular orbit about their common center of mass. The unit of mass is taken to be the total mass so that

$m_1 + m_2 = 1$

P1 has mass $1-\mu$, P2 has mass μ, $\mu = \dfrac{m_2}{m_1+m_2}$.

The equations of motion in the fixed $\bar{x}$-$\bar{y}$ bary-center system is

$$\ddot{x}_o = -\frac{k^2(1-\mu)(x_o - x_{o1})}{r_1{}^3} - \frac{k^2\mu(x_o - x_{o2})}{r_2{}^3}$$

$$\ddot{y} = -\frac{k^2(1-\mu)(y_o - y_{o1})}{r_1{}^3} - \frac{k^2\mu(y_o - y_{o2})}{r_2{}^3}$$

$$\ddot{z}_o = -\frac{k^2(1-\mu)z_o}{r_1{}^3} - \frac{k^2\mu z_o}{r_2{}^3}$$

where $r_1{}^2 = (x_{o1} - x_o)^2 + (y_{o1} - y_o)^2 + (z_{o1} - z_o)^2$ and $r_2{}^2 = (x_{o2} - x_o)^2 + (y_{o2} - y_o)^2 + (z_{o2} - z_o)^2$.

In a rotating system the equations of motion can be written as

$$\ddot{\bar{r}} + 2\bar{\omega} \times \bar{v} + \bar{\omega} \times (\bar{\omega} \times \bar{r}) = -\frac{k^2(1-\mu)}{r_1{}^3}\bar{r}_1 - \frac{k^2\mu}{r_2^3}\bar{r}_2 \qquad (7\text{-}1)$$

with

$\bar{\omega} = n\bar{k}$ $\qquad$ $\bar{r}_1 = (x - x_{o1})\bar{i} + y\bar{j} + z\bar{k}$

$\ddot{\bar{r}} = \ddot{x}\bar{i} + \ddot{y}\bar{j} + \ddot{z}\bar{k}$ $\qquad$ $\bar{r}_2 = (x - x_{o2})\bar{i} + y\bar{j} + z\bar{k}$

$\bar{v} = \dot{x}\bar{i} + \dot{y}\bar{j} + \dot{z}\bar{k}$

$\omega \times \bar{v} = -n(\dot{y}\bar{i} - \dot{x}\bar{j})$; $\bar{\omega} \times (\bar{\omega} \times \bar{r}) = -n^2(x\bar{i} + y\bar{j})$

where the x direction of the rotating system is such that the two massive bodies lie on the x axis and have coordinates $(-x_{o1}, 0, 0)$ and $(x_{o2}, 0, 0)$.

Hence the components in the rotating system become

$$\ddot{x} - 2n\dot{y} = n^2x - \frac{k^2(1-\mu)}{r_1{}^3}(x-x_{o1}) - \frac{k^2\mu}{r_2{}^3}(x-x_{o2})$$

$$\ddot{y} + 2n\dot{x} = n^2y - \frac{k^2(1-\mu)}{r_1{}^3}y - \frac{k^2\mu}{r_2{}^3}y \tag{7-2}$$

$$\ddot{z} = -\frac{k^2(1-\mu)}{r_1{}^3}z - \frac{k^2\mu}{r_2{}^3}z$$

where $(x_{o1}, 0)$ and $(x_{o2}, 0)$ are the coordinates of masses $(1-\mu)$ and μ respectively. Now let the distance between P1 and P2 $(x_{o2}-x_{o1})$ be unity and change the time scale so that $\tau = nt$. For a coordinate system centered at the center of mass, $x_{o1} = -\mu$ and $x_{o2} = 1-\mu$. From Kepler's "third" law

$$n = \frac{2\pi}{P} = \frac{\sqrt{m_1 + m_2}\; k}{a^{3/2}} . \tag{7-3}$$

But we have chosen the mass so that $m_1 + m_2 = 1$, the distance $a = 1$ and we choose time so that $n = 1$, i.e., $2\pi/P$ in τ units is 1, hence $k = 1$ by Kepler's law. The equations then become

$$\ddot{x} - 2\dot{y} = x - \frac{(1-\mu)(x+\mu)}{r_1{}^3} - \frac{\mu(x-1+\mu)}{r_2{}^3}$$

$$\ddot{y} + 2\dot{x} = y - \frac{(1-\mu)y}{r_1{}^3} - \frac{\mu}{r_2{}^3}y \tag{7-4}$$

$$\ddot{z} = -\frac{(1-\mu)z}{r_1{}^3} - \frac{\mu}{r_2{}^3}z$$

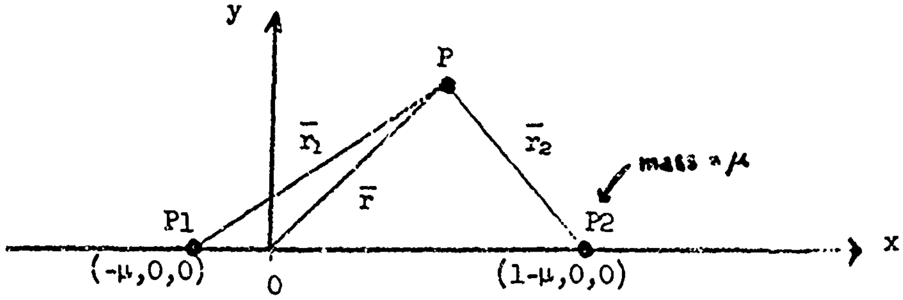


FIGURE 7-2

If we let

$$\tilde{U} = \frac{1}{2}(x^2 + y^2) + \frac{1-\mu}{r_1} + \frac{\mu}{r_2} \tag{7-5}$$

then since

$$r_1^2 = (x+\mu)^2 + y^2 + z^2 \tag{7-6}$$

$$r_2^2 = (x-1+\mu)^2 + y^2 + z^2$$

we can write

$$\frac{\partial\tilde{U}}{\partial x} = x - \frac{(1-\mu)(x+\mu)}{r_1^3} - \frac{\mu(x-1+\mu)}{r_2^3}$$

$$\frac{\partial\tilde{U}}{\partial y} = y - \frac{(1-\mu)y}{r_1^3} - \frac{\mu}{r_2^3}y \tag{7-7}$$

$$\frac{\partial\tilde{U}}{\partial z} = -\frac{(1-\mu)}{r_1^3}z - \frac{\mu}{r_2^3}z$$

With these our equations (7-4) can be written

$$\ddot{x} - 2\dot{y} = \frac{\partial\tilde{U}}{\partial x}$$

$$\ddot{y} + 2\dot{x} = \frac{\partial\tilde{U}}{\partial y} \tag{7-8}$$

$$\ddot{z} = \frac{\partial\tilde{U}}{\partial z} .$$

Multiplying these respectively by $2\dot{x}$, $2\dot{y}$, and $2\dot{z}$ and adding gives

$$2\dot{x}\ddot{x} + 2\dot{y}\ddot{y} + 2\dot{z}\ddot{z} = 2\dot{x}\,\frac{\partial\tilde{U}}{\partial x} + 2\dot{y}\,\frac{\partial\tilde{U}}{\partial y} + 2\dot{z}\,\frac{\partial\tilde{U}}{\partial z} \tag{7-9}$$

$$\frac{d}{dt}\,(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = 2\,\frac{d\tilde{U}}{dt} \tag{7-10}$$

$$d\tilde{U} = \frac{\partial\tilde{U}}{\partial x}\,dx + \frac{\partial\tilde{U}}{\partial y}\,dy + \frac{\partial\tilde{U}}{\partial z}\,dz.$$

Integrating (7-10) gives

$$v^2 = 2\tilde{U} - C \tag{7-11}$$

v is the magnitude of the velocity of P in the rotating coordinate system and C is the Jacobian constant.

Substituting for $\tilde{U}$ this Jacobi's integral gives

$$v^2 = (x^2 + y^2) + \frac{2(1-\mu)}{r_1} + \frac{2\mu}{r_2} - C\ . \tag{7-12}$$

In real motion v^2 must be positive. Thus for a given C value it is possible to draw contours in the x,y plane on which v = 0. These contours will be boundaries between regions where v^2 is positive and regions where (formally) v^2 is negative. A real particle cannot enter the region where v^2 is negative - hence cannot cross the contour $v^2 = 0$. These are the zero velocity curves - See Figure 7-3 on page 82. C can be determined by initial conditions on the particle. Note that once C is known, it is possible to compute particle speed, v, at any given point by Jacobi's equation. Its direction cannot be determined however, and thus the restricted three body problem remains unsolved in general.

Since angular momentum is not conserved, the motion is in general not planar. However, if the initial conditions are such that the z-components of position and velocity are zero, then the orbit will remain in the x-y plane.

In the rotating system the velocity is given by

$$\dot{\bar{r}}_o = \omega \times \bar{r} + \bar{v} \qquad \bar{\omega} = n\bar{k}$$

$$\bar{v} = \dot{x}\bar{i} + \dot{y}\bar{j} + \dot{z}\bar{k} \tag{7-13}$$

$$\dot{\bar{r}}_o = \bar{i}(\dot{x} - ny) + \bar{j}(\dot{y} + nx) + \bar{k}\dot{z}$$

The velocity with respect to τ then becomes (recall n=1, see page 78),

$$\dot{\bar{r}}_o = \bar{i}(\dot{x} - y) + \bar{j}(\dot{y} + x) + \bar{k}\dot{z} . \tag{7-14}$$

The kinetic energy of a particle of unit mass is

$$T = \frac{1}{2}\dot{\bar{r}}_o \cdot \dot{\bar{r}}_o = \frac{1}{2}(\dot{x}-y)^2 + \frac{1}{2}(\dot{y}+x)^2 + \frac{1}{2}\dot{z}^2 . \tag{7-15}$$

The potential energy is (in the inertial system),

$$V = -\frac{1-\mu}{r_1} - \frac{\mu}{r_2} . \tag{7-16}$$

The total energy is

$$2E = 2(T + V) = v^2 + 2(x\dot{y}-y\dot{x}) + (x^2+y^2) - \frac{2(1-\mu)}{r_1} - \frac{2\mu}{r_2} . \tag{7-17}$$

Since the potential energy in the rotating system is velocity dependent, the total energy E is not conserved, i.e., is not constant along a trajectory. We can, however, conserve something else, namely Jacobi's constant.

Substitute for v^2 from (7-12) into (7-17) to give

$$2E + C = 2(x^2 + y^2) + 2(x\dot{y} + y\dot{x}) = 2Q \qquad (7\text{-}18)$$

$$Q = (x^2 + y^2) + x\dot{y} - y\dot{x} \quad . \qquad (7\text{-}19)$$

In polar coordinates $x = r \cos f \quad y = r \sin f$ so--

$$Q = r^2 + r^2\dot{f} = r^2 \left[1 + \frac{df}{d\tau}\right] = r^2 \left[1 + \frac{1}{n}\ \frac{df}{dt}\right]$$

$$nQ = r^2\ (n + \dot{f}) = r^2\ \dot{\phi} = h_z \qquad (7\text{-}20)$$

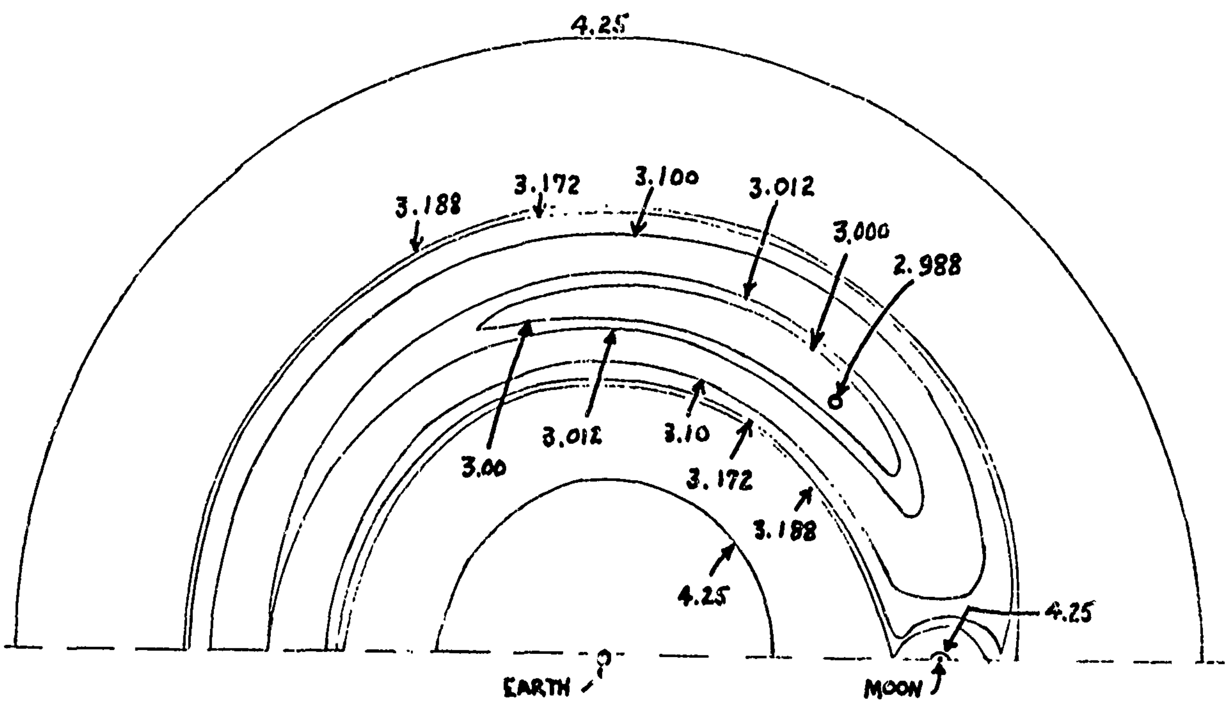


FIGURE 7-3 CONTOURS OF CONSTANT POTENTIAL ENERGY IN ROTATING SYSTEM WITH Z = 0 AND μ= 0.01213 (Coordinates at Center of Mass)

where h_z is the angular momentum of a unit mass about the z axis (or Z_0 axis). We thus have

$$-C = 2(E - h_z)$$

therefore the quantity $E-h_z$ is conserved over any given trajectory.

We have been able to obtain this extra integral because we found Jacobi's integral which in turn was possible because the two large masses move in a circular orbit. If the orbit is elliptical about their common center of mass, we can no longer obtain this integral.

This restricted three body problem forms the first order model used by Hill to obtain the motion of the moon. Hill's work was used by Brown to compute the actual motion.

These zero velocity curves (Figure 7-3) deserve further mention. We can use them to delineate regions of possible motion.

Consider the zero velocity curves shown in Figure 7-3. If both C and x^2+y^2 are large, then by equation (7-12) $x^2+y^2 \approx C_1$ is the equation of a circle. In Figure 7-3 this is the case for C = 4.25. However, C can also be large ($C = C_1$) if either r_1 or r_2 is very small and, hence for large $C = C_1$ we have the case below. The z axis is perpendicular to the plane of the paper.

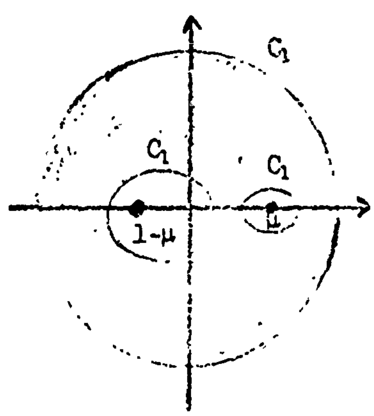


FIGURE 7-4

The shaded area of Figure 7-4 represents the volume of space where the particle's velocity would be imaginary and therefore inaccessible. If the particle starts off inside one of the ovals or outside the larger circle, the particle must remain within its initial domain since the three regions are separated by the "forbidden" domain.

As C then becomes small, the inner ovals expand while the outer surface of almost circular cross-section shrinks. For a certain value of C, say C_2, the inner ovals meet at the libration point L1. This is illustrated in Figure 7-5.

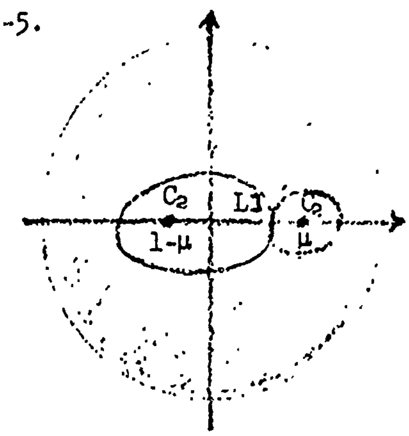


FIGURE 7-5

A slight decrease in C now results in the ovals coalescing to form a dumbbell-shaped surface with a narrow neck. The particle is then free to travel between the massive bodies. This is shown in Figure 7-6(a). For a further decrease, the inner region meets the outer at the L2 libration point as shown in Figure 7-6(b).

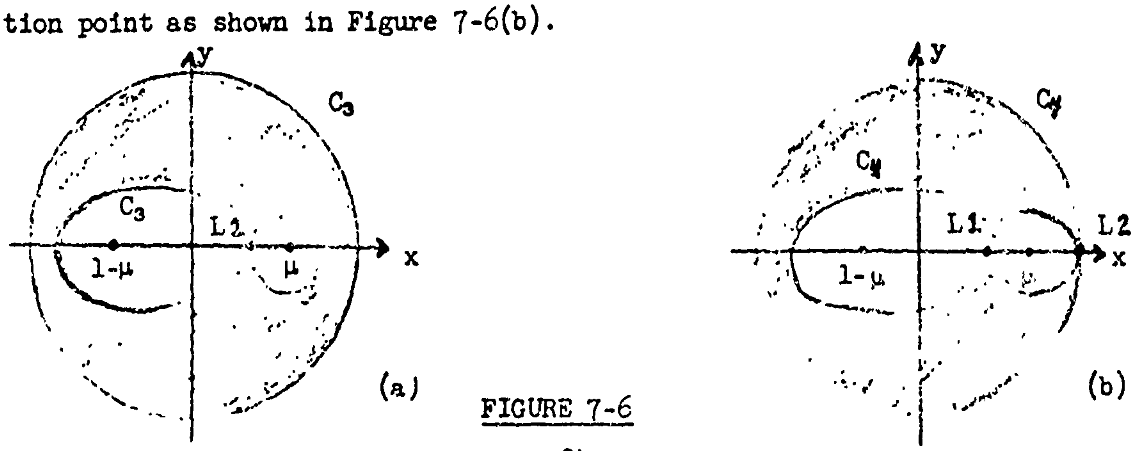


FIGURE 7-6

As C is decreased still further, the forbidden region shrinks at the L3 libration point on the left and opens about the L2 libration point thus allowing the particle to wonder out of the region of the two finite masses and into outer space. As the process continues, the regions inaccessible to the particle in the x-y plane shrink until they vanish at the final two libration points L4 and L5.

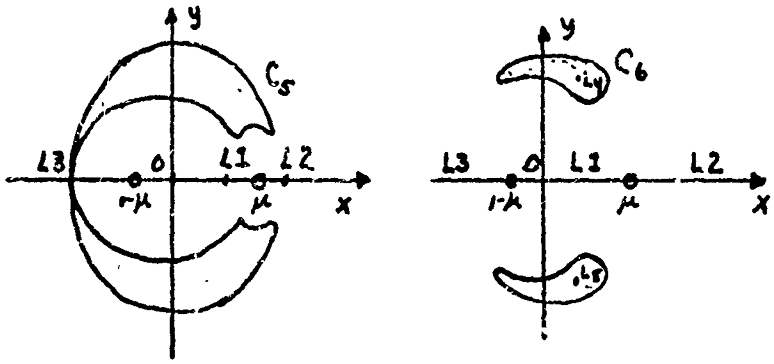


FIGURE 7-7

The behavior of the zero velocity surfaces as C change in the xz and yz planes are shown on the next page where the corresponding values of C are shown.

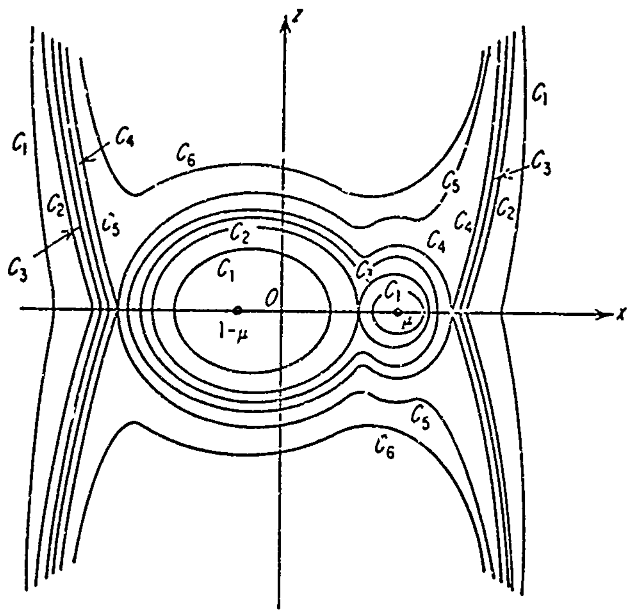


Figure 7-8

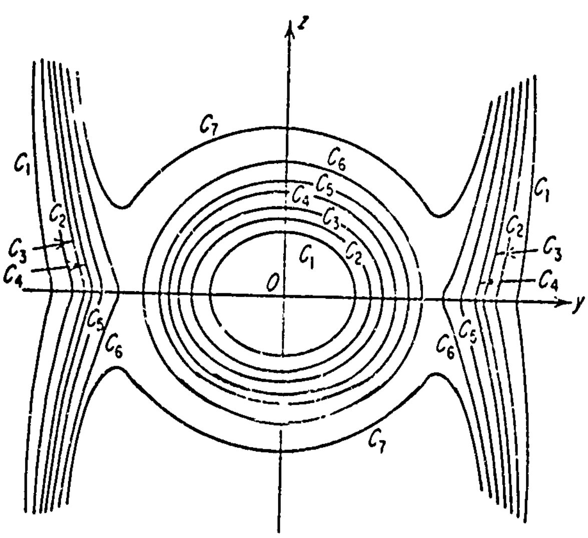


Figure 7-9

8. REGULARIZATION IN 3-BODY PROBLEM

One of the uses of the restricted three-body problem is to calculate trajectories from earth to moon. These essentially require a collision. As the moon is approached, the particle velocity increases in a very short time. This requires extremely small time intervals in any numerical integration of the orbit. Our problem is one caused by a singularity of the differential equation of motion. The idea of Levi-Civita and Sundman was to introduce a new independent variable, a pseudo time (τ) instead of the ordinary time (t) in such a manner that τ would change uniformly during collision. The new and old time differentials are usually related by

$$d\tau = \frac{dt}{f(r)} \tag{8-1}$$

where $f(r)$ is a function of the distance between the colliding bodies. This removes the singularity and such a process is called regularization.

To see the essentials consider a simplified two-body problem with the mass of one being unity and the other mass very very small. The equation of motion of the small mass is

$$\ddot{\bar{r}} = -\frac{\bar{r}}{|\bar{r}|^3} \tag{8-2}$$

where $\bar{r}$ is the relative position vector of the small mass and the units of mass and time have been chosen to make $k^2\ (m_1 + m_2) = 1.0$. If the

small mass is given initial conditions such that its motion occurs along a straight line through the larger mass (along a rectilinear orbit) the vector character of $\bar{r}$ can be ignored and we have $\ddot{r} = -\frac{1}{r^2}$ where r is now the scalar distance between the bodies. The singularity of this equation is at $r = 0$, i.e., at the point where the moving body and larger, fixed body collide.

Now integrate this equation to give

$$\dot{r}^2 = \frac{2}{r} - C \,. \tag{8-3}$$

At the singularity ($r = 0$) the acceleration and velocity are infinite. In fact, close to $r = 0$, we have

$$\frac{dr}{dt} = \pm\sqrt{\frac{2}{r}} \tag{8-4}$$

which says the velocity approaches infinity like $\frac{1}{\sqrt{r}}$ as $r \to 0$.

The Levi-Civita idea of regularization is to introduce τ so that

$$d\tau = \frac{dt}{f(r)} = \frac{dt}{r} \,. \tag{8-5}$$

Our velocity then can be written

$$\frac{dr}{d\tau} = \pm\frac{dt}{d\tau}\sqrt{\frac{2}{r}} = \pm\sqrt{2r}\,, \quad \frac{dt}{d\tau} = r \tag{8-6}$$

The new velocity $\frac{dr}{d\tau}$ is now zero at the previous singularity ($r = 0$) and the actual velocity is regular, i.e.,

$$\frac{dr}{d\tau} = \sqrt{r\,(2 - Cr)} \tag{8-7}$$

and the original second-order differential equation becomes

$$\frac{d^2r}{d\tau^2} + Cr = 1 \tag{8-8}$$

which is also devoid of singularities. However, with $f(r) = r$ our new equation gives approximately $\Delta r \approx \pm\, \Delta\tau \sqrt{2r}$ so that as $r \to 0$ the integration step size ($\Delta\tau$) must be increased to maintain a constant Δr. Note that for the regular time we have $\Delta r = \pm\, \Delta t \sqrt{\frac{2}{r}}$ and as $r \to 0$, Δt has to be decreased in order to have a constant Δr.

Now with a slight generalization we can improve the problem even further. Let

$$d\tau = \frac{dt}{\sqrt{r}} \tag{8-9}$$

then close to collision we now have the relation

$$\frac{dr}{d\tau} = \pm\sqrt{2} \tag{8-10}$$

and the complete expression for the pseudo velocity becomes

$$\frac{dr}{d\tau} = \pm\, \sqrt{2 - Cr}\ . \tag{8-11}$$

The velocity at the singularity is $\pm\sqrt{2}$, and, as $r \to 0$ we now have $\Delta r \approx \sqrt{2}\,\Delta\tau$, i.e., integration in this pseudo system uses a constant step size near collision. The original differential equation now becomes

$$\frac{d^2 r}{d\tau^2} = -C \tag{8-12}$$

which is not only regular but also considerably simpler than either the other two representations.

For the case of the restricted three-body problem we have the set of equations

$$\ddot{x} = x + 2\dot{y} - \frac{(1-\mu)(x+\mu)}{r_1^{\,3}} - \mu\,\frac{x+\mu-1}{r_2^{\,3}} \tag{8-13}$$

$$\ddot{y} = y - 2\dot{x} - \frac{(1-\mu)\,y}{r_1^{\,3}} - \frac{\mu y}{r_2^{\,3}} \;.$$

The Thiele transformation will regularize the equations by using the following transformations:

$$x = \frac{1}{2} - \mu + \frac{1}{2}\cos u \cosh v$$

$$y = -\frac{1}{2}\sin u \sinh v \;, \tag{8-14}$$

The new equations are then functions of u and v. This transformation takes as a new independent variable the variable s defined by

$$s = \int_0^t \frac{dt}{r_1\, r_2} \;. \tag{8-15}$$

This eliminates the need to vary the integration step size as the trajectory passes near either body. Making these substitutions into the rotating frame results in the following equations:

$$\frac{d^2u}{ds^2} = \frac{1}{2}\left(\cosh^2 v - \cos^2 u\right)\frac{dv}{ds}$$

$$+ \sin u\left[\frac{\sigma}{2} + \left(C + \frac{\sigma^2}{16}\right)\cos u - \frac{\sigma}{16}\cosh v\left(\cosh^2 v - 3\cos^2 u\right)\right.$$

$$\left. + \frac{1}{16}\cos u\cos 2u\right] \tag{8-16}$$

$$\frac{d^2v}{ds^2} = -\frac{1}{2}\left(\cosh^2 v - \cos^2 u\right)\frac{du}{ds}$$

$$+ \sinh v\left[\frac{1}{2} + \left(C + \frac{\sigma^2}{16}\right)\cosh v - \frac{\sigma}{16}\cos u\left(\cos^2 u - 3\cosh^2 v\right)\right.$$

$$\left. + \frac{1}{16}\cosh u\cosh 2v\right]$$

where $\sigma = 1 - 2\mu$ and C is the Jacobi constant which is defined by the following:

$$\left(\frac{du}{ds}\right)^2 + \left(\frac{dv}{ds}\right)^2 + \sigma\cos u - \cosh v - C\left(\cosh^2 v - \cos^2 u\right) \tag{8-17}$$

$$- \frac{1}{16}\sin^2 u\left(\cos u + \sigma\cosh v\right)^2 - \frac{1}{16}\sinh^2 v\left(\sigma\cos u + \cosh v\right)^2 = 0$$

One can plot the motion in a fixed inertial frame of reference by using

$$t = \frac{1}{4}\int_0^s (\cosh^2 v - \cos^2 u)\, ds \tag{8-18}$$

to generate a new variable from which one computes

$$z_1 = x \cos t - y \sin t$$

$$z_2 = y \cos t + x \sin t \,.$$

In spite of its messy appearance the terms in the Thiele equations are better behaved in position and velocity terms. The sinusoidal terms are bounded between +1 and -1 and the hyperbolic functions do not vary greatly in range. There are no singularities.

Given the initial conditions in the rotating coordinate system of

$$\mu,\ x(o),\ y(o),\ \dot{x}(o),\ \dot{y}(o),$$

One then proceeds as follows:

$$r_1(o) = \sqrt{(x(o) + \mu)^2 + y^2(o)} \quad ; \quad r_2(o) = \sqrt{(x(o) + \mu - 1)^2 + y^2(o)} \,.$$

$$\cosh v(o) = r_1(o) + r_2(o) \tag{8-20}$$

$$\cos u(o) = r_1(o) - r_2(o)$$

$$(\cosh v(o))^2 = (r_1(o) + r_2(o))^2 \qquad (\cos u(o))^2 = (r_1(o) - r_2(o))^2$$

$$\sinh v(o) = \sqrt{\cosh^2 v(o) - 1} \qquad \sin u(o) = \sqrt{1 - \cos^2 u(o)}$$

$$\sigma = 1 - 2\mu \qquad \sigma^2 = (1 - 2\mu)^2$$

$$\frac{du(o)}{ds} = \frac{1}{2}\left[- \dot{y}(o) \cos u(o) \sinh v(o) - \dot{x}(o) \sin u(o) \cosh v(o)\right]$$

$$\frac{dv(o)}{ds} = \frac{1}{2}\left[- \dot{y}(o) \sin u(o) \cosh v(o) + \dot{x}(o) \cos u(o) \sinh v(o)\right]$$

$$C = \frac{1}{[\cosh^2 v(o) - \cos^2 u(o)]} \left\{\left(\frac{dv(o)}{ds}\right)^2 + \left(\frac{du(o)}{ds}\right)^2 + \sigma \cos u(o)\right.$$

$$- \cosh v(o) - \frac{1}{16} (\sigma \cos u(o) + \cosh v(o))^2 \sinh^2 v(o)$$

$$\left. - \frac{1}{16} (\cos u(o) + \sigma \cosh v(o))^2 \sin^2 u(o)\right\} .$$

One then has all the constants and initial conditions to proceed with the numerical integration of the Thiele Equations (8-16). The results in u and v are transformed back to x and y by

$$x = \frac{1}{2} - \mu + \frac{1}{2} \cos u \cosh v$$
$$y = - \frac{1}{2} \sin u \sinh v \tag{8-21}$$

9. STABILITY OF TRAJECTORIES

For the restricted 3-body problem we have

$$\left.\begin{aligned} \ddot{x} - 2\dot{y} &= \tilde{U}_x = \frac{\partial \tilde{U}}{\partial x} \\ \ddot{y} + 2\dot{x} &= \tilde{U}_y = \frac{\partial \tilde{U}}{\partial y} \\ \ddot{z} &= \tilde{U}_z = \frac{\partial \tilde{U}}{\partial z} \end{aligned}\right\} \tag{9-0}$$

Massive bodies at $(x_1, 0, 0)$ and $(x_2, 0, 0)$

Rotating barycenter coordinates.

Planar motion.

$$\left.\begin{aligned} \tilde{U}_x &= x - \frac{(1-\mu)(x - x_1)}{r_1^3} - \frac{\mu}{r_2^3}(x - x_2) \\ \tilde{U}_y &= y - \frac{(1-\mu)\,y}{r_1^3} - \frac{\mu y}{r_2^3} \\ \tilde{U}_z &= -\frac{(1-\mu)\,z}{r_1^3} - \frac{\mu z}{r_2^3} \end{aligned}\right\} \tag{9-1}$$

$$\left.\begin{aligned} r_1 &= \left[(x - x_1)^2 + y^2 + z^2\right]^{1/2} \\ r_2 &= \left[(x - x_2)^2 + y^2 + z^2\right]^{1/2} \end{aligned}\right\} \tag{9-2}$$

To study stability about one of the Lagrangian points we try to determine if a very small displacement from that point will result in a bounded orbit near the point. Let the coordinates of one of the libration points be (x_o, y_o, z_o). Let α, β, γ denote small displacements of a particle from this (x_o, y_o, z_o) point so that

$$
\begin{array}{lll}
x = x_o + \alpha & \dot{x} = \dot{x}_o + \dot{\alpha} & \ddot{x} = \ddot{x}_o + \ddot{\alpha} \\
y = y_o + \beta & \dot{y} = \dot{y}_o + \dot{\beta} & \ddot{y} = \ddot{y}_o + \ddot{\beta} \\
z = z_o + \gamma & \dot{z} = \dot{z}_o + \dot{\gamma} & \ddot{z} = \ddot{z}_o + \ddot{\gamma} .
\end{array}
\tag{9-3}
$$

Since the displacements are small we can expand the $\tilde{U}_x$, $\tilde{U}_y$, and $\tilde{U}_z$ terms in a Taylor's series about the point (x_o, y_o, z_o); thus - - -

$$
\begin{aligned}
\tilde{U}_x &= (\tilde{U}_x)_o + \alpha\, (\tilde{U}_{xx})_o + \beta\, (\tilde{U}_{xy})_o + \gamma\, (\tilde{U}_{xz})_o \\
\tilde{U}_y &= (\tilde{U}_y)_o + \alpha\, (\tilde{U}_{yx})_o + \beta\, (\tilde{U}_{yy})_o + \gamma\, (\tilde{U}_{yz})_o \\
\tilde{U}_z &= (\tilde{U}_z)_o + \alpha\, (\tilde{U}_{zx})_o + \beta\, (\tilde{U}_{zy})_o + \gamma\, (\tilde{U}_{zz})_o
\end{aligned}
\tag{9-4}
$$

The zero subscript means evaluation at the point (x_o, y_o, z_o). At this equilibrium point we found before that

$$
\begin{aligned}
&(\tilde{U}_x)_o = (\tilde{U}_y)_o = (\tilde{U}_z)_o = 0 \\
&\ddot{x}_o = \ddot{y}_o = \ddot{z}_o = \dot{x}_o = \dot{y}_o = \dot{z}_o = 0
\end{aligned}
\tag{9-5}
$$

Indeed this is the definition of an equilibrium point. Using the Taylor series expansion and substituting for x, y, z gives

$$
\begin{aligned}
\ddot{\alpha} - 2\dot{\beta} &= \alpha\, (\tilde{U}_{xx})_o + \beta\, (\tilde{U}_{xy})_o + \gamma\, (\tilde{U}_{xz})_o \\
\ddot{\beta} + 2\dot{\alpha} &= \alpha\, (\tilde{U}_{yx})_o + \beta\, (\tilde{U}_{yy})_o + \gamma\, (\tilde{U}_{yz})_o \\
\ddot{\gamma} &= \alpha\, (\tilde{U}_{zx})_o + \beta\, (\tilde{U}_{zy})_o + \gamma\, (\tilde{U}_{zz})_o .
\end{aligned}
\tag{9-6}
$$

These are the linearized equations of motion for a particle about the libration point. The Taylor series coefficients are found from the following:

$$
\begin{aligned}
\tilde{U}_{xx} &= 1 - \frac{1-\mu}{r_1{}^3} - \frac{\mu}{r_2{}^3} + \frac{3\,(1-\mu)(x-x_1)^2}{r_1{}^5} + \frac{3\mu\,(x-x_2)^2}{r_2{}^5} \\
\tilde{U}_{xy} &= \frac{3\,(1-\mu)(x-x_1)\,y}{r_1{}^5} + \frac{3\mu\,(x-x_2)\,y}{r_2{}^5} \\
\tilde{U}_{xz} &= \frac{3\,(1-\mu)(x-x_1)\,z}{r_1{}^5} + \frac{3\mu\,(x-x_2)\,z}{r_2{}^5} \\
\tilde{U}_{yy} &= 1 - \frac{1-\mu}{r_1{}^3} - \frac{\mu}{r_2{}^3} + \frac{3\,(1-\mu)\,y^2}{r_1{}^5} + \frac{3\mu y^2}{r_2{}^5} \\
\tilde{U}_{yz} &= \frac{3\,(1-\mu)\,yz}{r_1{}^5} + \frac{3\mu yz}{r_2{}^5} \\
\tilde{U}_{zz} &= -\frac{1-\mu}{r_1{}^3} - \frac{\mu}{r_2{}^3} + \frac{3\,(1-\mu)\,z^2}{r_1{}^5} + \frac{3\mu z^2}{r_2{}^5}
\end{aligned}
\tag{9-7}
$$

We now ask if equations (9-6) are stable. First consider the libration point at L_4.

$$
\begin{aligned}
r_1^2 &= r_2^2 = 1 \\
(x_o + \mu)^2 + y_o{}^2 &= (x_o - 1 + \mu)^2 + y_o{}^2 = 1
\end{aligned}
\tag{9-8}
$$

Therefore

$$x_o = \frac{1}{2} - \mu$$

$y_o{}^2 = \frac{3}{4}$ $\qquad y_o = \frac{\sqrt{3}}{2}$ $\qquad z_o = 0$

$x_o - x_1 = \frac{1}{2}$ $\qquad x_o - x_2 = -\frac{1}{2}$ $\qquad x_2 - x_o = \frac{1}{2}$

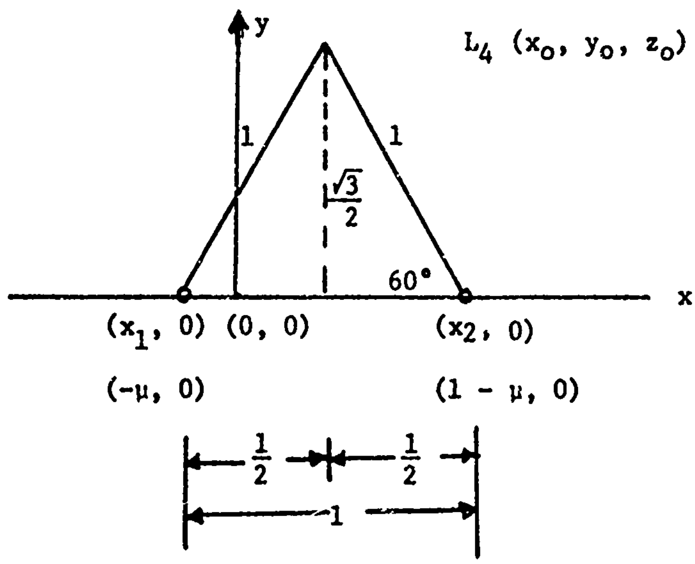


FIGURE 9-1

Evaluation of the partials at $(x_o, y_o, z_o) = (\frac{1}{2} - \mu, \frac{\sqrt{3}}{2}, 0)$ gives the following results:

$$\tilde{U}_{xx} = \frac{3}{4} \qquad \tilde{U}_{yz} = \tilde{U}_{zy} = 0$$

$$\tilde{U}_{xy} = \tilde{U}_{yx} = \frac{3\sqrt{3}}{4}(1 - 2\mu) \qquad \tilde{U}_{yy} = \frac{9}{4} \tag{9-9}$$

$$\tilde{U}_{yz} = \tilde{U}_{zx} = 0 \qquad \tilde{U}_{zz} = -1$$

Thence Equation (9-6) becomes

$$\ddot{\alpha} - 2\dot{\beta} = \frac{3}{4}\alpha + \frac{3\sqrt{3}}{4}(1 - 2\mu)\beta$$

$$\ddot{\beta} + 2\dot{\alpha} = \frac{3\sqrt{3}}{4}(1 - 2\mu)\alpha + \frac{9}{4}\beta \tag{9-10}$$

$$\ddot{\gamma} = -\gamma ,$$

This last equation has a solution $\gamma = C_1 \cos t + C_2 \sin t$ and hence motion in γ is bounded. We could solve the other two equations by the LaPlace transform method and thereby show the motion is bounded, but let us try the classical approach. Assume a solution of the form

$$\alpha = Ae^{\lambda t} \qquad \beta = Be^{\lambda t} \tag{9-11}$$

Substituting this in Equation (9-10) gives

$$A\lambda^2 e^{\lambda t} - 2B\lambda e^{\lambda t} = \frac{3}{4}Ae^{\lambda t} + \frac{3\sqrt{3}}{4}(1 - 2\mu)Be^{\lambda t}$$

$$B\lambda^2 e^{\lambda t} + 2A\lambda e^{\lambda t} = \frac{3\sqrt{3}}{4}(1 - 2\mu)Ae^{\lambda t} + \frac{9}{4}Be^{\lambda t} \tag{9-12}$$

which can be written as

$$A\left[\lambda^2 - \frac{3}{4}\right] + B\left[-2\lambda - \frac{3\sqrt{3}}{4}(1 - 2\mu)\right] = 0$$

$$A\left[2\lambda - \frac{3\sqrt{3}}{4}(1 - 2\mu)\right] + B\left[\lambda^2 - \frac{9}{4}\right] = 0 \tag{9-13}$$

There will be a nontrivial solution for any A and B if

$$\begin{vmatrix} \lambda^2 - \frac{3}{4} & -2\lambda - \frac{3\sqrt{3}}{4}(1 - 2\mu) \\ 2\lambda - \frac{3\sqrt{3}}{4}(1 - 2\mu) & \lambda^2 - \frac{9}{4} \end{vmatrix} = 0 \qquad (9\text{-}14)$$

or

$$4\lambda^4 + 4\lambda^2 + 27\mu\,(1 - \mu) = 0\,. \qquad (9\text{-}15)$$

The solution is

$$\lambda^2 = -\frac{1}{2} \pm \frac{1}{2}\sqrt{1 - 27\mu\,(1 - \mu)}$$

To have bounded motion, λ must be imaginary so $\lambda^2 < 0$. For this to hold $1 - 27\mu\,(1 - \mu) \geq 0$

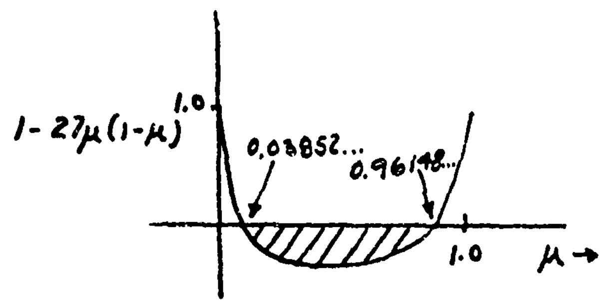


Figure 9-2

$$0 \leq \mu < 0.0385207\ .\ .\ . \qquad (9\text{-}16)$$

The other root gives $0.96148 \ldots \leq \mu < 1$

Since

$$\mu = \quad = \frac{m_2}{m_1 + m_2}$$

then the above stability requires that

$$\frac{m_2}{m_1} \quad < 0.040064 \ldots$$

Note that this is the stability of the linearized equation. Stability in the large has not been proven here. For that we need more powerful tools such as the second method of Liapunov.

The technique above may be used at the other libration points. When this is done one discovers that the L1, L2, L3 libration points are unstable for all mass ratios. See Moulton, P 300, or McCuskey, P 118.

REFERENCES

1. Roy, A. E., *The Foundations of Astrodynamics*, Macmillan (1965).

2. Danby, J. M. A., *Fundamentals of Celestial Mechanics*, Macmillan (1962).

3. Brouwer, D. and C. M. Clemence, *Methods of Celestial Mechanics*, Academic Press (1961).

4. Moulton, F. R., *Celestial Mechanics*, Macmillan (1914).

5. McCuskey, S. W., *Introduction to Celestial Mechanics*, Addison-Wesley (1963).

6. Szebehely, V., "The Restricted Problems of Three Bodies," Notes of the Summer Institute in Dynamical Astronomy, Yale University (August 1961), See also his forthcoming book, *Theory of Orbits* by Academic Press.

7. Deutsch, R., *Orbital Dynamics of Space Vehicles*, Prentice-Hall (1963).

8. Szebehely, V., D. A. Pierce and E. Myles Standish, Jr., "A Group of Earth-to-Moon Trajectories with Consecutive Collesions" in Celestial Mechanics and Astrodynamics edited by V. Szebehely, Vol. 14 of Progress in Astronautics and Aeronautics, Academic Press (1964).

10. HAMILTON-JACOBI THEORY

Canonical Systems

For a particle of mass m, located at x_1, x_2, x_3 we have

$$m \frac{d^2 x_i}{dt^2} = - \frac{\partial V}{\partial x_i} \qquad i = 1,2,3, \tag{10-1}$$

Where $V(x_1, x_2, x_3, t)$ is a potential. These equations are three of second order. We can express these as six equations of first order. To do so we introduce three new variables, y_1, y_2, y_3 called momenta by defining

$$y_i = m \frac{dx_i}{dt} \qquad i = 1,2,3. \tag{10-2}$$

The kinetic energy of the system is

$$T = \frac{1}{2}\sum m \left(\frac{dx_i}{dt}\right)^2 = \frac{1}{2}\sum \frac{1}{m} y_i^2 . \tag{10-3}$$

Differentiate (10-2) and substitute into (10-1) to give

$$\frac{dy_i}{dt} = - \frac{\partial V}{\partial x_i} \qquad i = 1,2,3. \tag{10-4}$$

From (10-2) and (10-3) we have

$$\frac{dx_i}{dt} = \frac{\partial T}{\partial y_i} \qquad i = 1,2,3. \tag{10-5}$$

Since V and T are independent of y_i and x_i respectively, (10-4) and (10-5) may be written, with

$$H = T + V = T(y_i) + V(x_i)$$

as

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial y_i}; \quad \frac{dy_i}{dt} = -\frac{\partial H}{\partial x_i}, \qquad i = 1,2,3. \tag{10-6}$$

This set of differential equations is called the canonical or Hamiltonian form and H is called the Hamiltonian function. Actually we have a special case of a more general formulation. Before discussing this, note that we can extend the equations to 2n variables. These will satisfy the canonical equations.

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial y_i}; \quad \frac{dy_i}{dt} = -\frac{\partial H}{\partial x_i} \qquad i = 1,2\ldots n \tag{10-7}$$

where

$H = H(x_i, y_i)$ i.e., it does not contain any $\dot{x}_i$ terms.

We can also derive these equations in more general form from Lagrange's equations. If a conservative holonomic dynamical system*

*See the discussion in Appendix II (page 21) or Goldstein, page 14, et. seq.

can be described by the generalized coordinates q_i, $i = 1,2,\ldots,n$, then the Lagrange equations of motion are

$$\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}_i}\right] - \frac{\partial L}{\partial q_i} = 0 \tag{10-8}$$

where the Lagrangian potential is $L(q_i, \dot{q}_i, t)$ with $L = T - V$.

$$L = T - V \tag{10-9}$$

$T(\dot{q}_i, q_i, t)$ is the kinetic energy of the system, and $V(q_i, t)$ is the potential energy.

These equations are of order 2n and thus we require 2n state variables to represent the system. To do this we select n variables q_i and n variables p_i, which we arbitrarily define as

$$p_i = \frac{\partial L}{\partial \dot{q}_i}\,. \tag{10-10}$$

These p_i are called the generalized momenta and basically replace the $\dot{q}_i$ variables. The original Lagrange equation is in terms of q_i and $\dot{q}_i$; we are now transforming to q_i, p_i system. This is done by means of a Legendre transformation (see Goldstein, page 215).

Consider the so-called Hamiltonian function

$$H(p, q, t) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t). \qquad (10\text{-}11)$$

Assuming that H has continuous first partial derivatives, its total differential is

$$dH = \sum_i \frac{\partial H}{\partial p_i} dp_i + \sum_i \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial t} dt. \qquad (10\text{-}12)$$

Differentiation of (10-11) gives,

$$dH = \sum_i \dot{q}_i dp_i + \sum_i \left(p_i - \frac{\partial L}{\partial \dot{q}_i}\right) d\dot{q}_i - \sum \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt \qquad (10\text{-}13)$$

By definition, $p_i = \frac{\partial L}{\partial \dot{q}_i}$, so the second term vanishes. From Lagrange's equation

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = \frac{dp_i}{dt} = \frac{\partial L}{\partial q_i} \quad , \quad i = 1, 2, \ldots n.$$

so that (10-13) becomes

$$dH = \sum_i \frac{dq_i}{dt} dp_i - \sum_i \frac{dp_i}{dt} dq_i - \frac{\partial L}{\partial t} dt \qquad (10\text{-}14)$$

Comparing this with (10-12) shows that

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \qquad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \qquad (10\text{-}15)$$

We thus again arrive at the canonical equations. The first two relations are known as <u>Hamilton's canonical equations</u>. They constitute a set of 2n first-order differential equations for the quant'ties q_i, p_i, $i = 1, \ldots n$. These equations describe the motion of the system just as does the set of n second-order Lagrange equations.

No phys'cal principles have been employed in this development, other than those inherent in Lagrange's equation. Since Lagrange's equation may be derived from either Newton's second law or Hamilton's principle (Goldstein, page 225), the canonical equations have the same basis.

This being the case, one might wonder why the extra trouble for Hamilton. The principal reason is convenience in analytical manipulations which will become quite clear as the course proceeds.

If one is interested only in deriving the equations of motion, the canonical equations are usually more awkward than Lagrange's equations. We desire much more, however, and the symmetry of these equations will become quite helpful.

The process of finding $H(p_i, q_i, t)$ then becomes as follows:

1. Calculate T, V and thence $L = T - V$
2. Determine the generalized momenta from $p_i = \frac{\partial L}{\partial \dot{q}_i}$.
3. Form

$$H = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i, t)$$

and eliminate all $\dot{q}_i$ by replacing them by equivalent expressions involving p_i, q_i by using $p_i = \frac{\partial L}{\partial \dot{q}_i}$.

The variable $\dot{q}_i$ usually appears only in the kinetic energy term and there only in quadratic form, hence $p_i = \frac{\partial L}{\partial \dot{q}_i}$ is nearly always a linear function of $\dot{q}_i$ and we can easily solve for $\dot{q}_i$ in terms of p_i and q_i. Thus we are able to form $H(p_i, q_i, t)$. To illustrate this advantage assume we choose a different set of variables $Y(\dot{p}_i, \dot{q}_i, t)$. We define

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad \text{and hence} \quad \frac{d}{dt}(p_i) = \frac{\partial L}{\partial q_i} \tag{10-16}$$

then defining

$$Y = \sum_i \dot{p}_i q_i - L(q_i, \dot{q}_i, t) \tag{10-17}$$

which we could call the Yamiltonian, we can proceed as with H to derive canonical equations.

$$dY = \sum_i (\dot{p}_i dq_i + q_i d\dot{p}_i - \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt) \tag{10-18}$$

Using equation (10-16) this becomes

$$dY = \sum_i (\dot{p}_i dq_i + q_i d\dot{p}_i - \dot{p}_i dq_i - p_i d\dot{q}_i - \frac{\partial L}{\partial t} dt) \tag{10-19}$$

$$dY = \sum_i \left(q_i \, d\dot{p}_i - p_i \, d\dot{q}_i - \frac{\partial L}{\partial t} dt \right) \tag{10-20}$$

But since $Y = Y(\dot{p}_i, \dot{q}_i, t)$ we have

$$dY = \sum_i \left(\frac{\partial Y}{\partial \dot{p}_i} d\dot{p}_i + \frac{\partial Y}{\partial \dot{q}_i} d\dot{q}_i + \frac{\partial Y}{\partial t} dt \right) \tag{10-21}$$

and comparison of (10-20) with (10-21) shows that

$$q_i = \frac{\partial Y}{\partial \dot{p}_i} \quad \text{and} \quad p_i = -\frac{\partial Y}{\partial \dot{q}_i} \quad , \quad \frac{\partial L}{\partial t} = -\frac{\partial Y}{\partial t} \, , \tag{10-22}$$

and the problem appears to be neatly solved. However, in order to find $Y(\dot{p}_i, \dot{q}_i, t)$ we must solve equation (10-16) for $q_i = q_i\ (\dot{p}_i, \dot{q}_i, t)$ and substitute into (10-17) to obtain $Y(\dot{p}_i, \dot{q}_i, t)$. Since q_i does not usually appear linearly in the Lagrangian L, equation (10-16) cannot easily be solved as $q_i = q_i\ (\dot{p}_i, \dot{q}_i, t)$. Thus, alas, we abandon the Yamiltonian for more difficult but greener pastures.

11. SIGNIFICANCE OF THE HAMILTONIAN FUNCTION

In addition to the advantage in mathematical manipulation, the Hamiltonian also has a physical significance. If the Lagrangian does not contain time explicitly then $\frac{\partial L}{\partial t} = 0$. In this case the Hamiltonian will not contain time explicitly either so $\frac{\partial H}{\partial t} = 0$.

Differentiating H we have,

$$\frac{dH(p_i, q_i, t)}{dt} = \sum_i \frac{\partial H}{\partial p_i} \dot{p}_i + \sum \frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial t} \tag{11-1}$$

Making use of the canonical equations,

$$\dot{H} = \sum_i \left(- \frac{\partial H}{\partial p_i} \frac{\partial H}{\partial q_i} + \frac{\partial H}{\partial q_i} \frac{\partial H}{\partial p_i} \right) + \frac{\partial H}{\partial t} . \tag{11-2}$$

The first term vanishes identically, so

$$\dot{H} = \frac{\partial H}{\partial t} = 0 \tag{11-3}$$

If the Lagrangian (and hence the Hamiltonian) does not contain t explicitly then $\frac{\partial H}{\partial t} = 0$, and H is a constant of the motion.

Under a different set of circumstances the Hamiltonian is equal to the total energy (see the last two pages of Appendix II).

If the kinetic energy can be expressed as a (positive definite) homogeneous quadratic form in the $\dot{q}_i$, as,

$$T = \sum_{k,\ell} a_{k\ell}(q,\ t)\ \dot{q}_k\dot{q}_\ell \qquad k = 1,2,\ldots n \quad \ell = 1,2,\ldots n \tag{11-4}$$

with

$$a_{k\ell} = a_{\ell k} \quad \ell \neq k\ .$$

The Lagrangian may be written

$$L = \sum_{k,\ell} a_{k\ell}(q,\ t)\ \dot{q}_k\dot{q}_\ell - V(q,t) \tag{11-5}$$

where V is the potential energy. The generalized momenta are

$$p_i = \frac{\partial L}{\partial \dot{q}_i} = 2\sum_k a_{ki}(q,t)\dot{q}_k. \tag{11-6}$$

The Hamiltonian is, then:

$$H = 2\sum_{\ell,k} a_{\ell k}(q,t)\dot{q}_k\dot{q}_\ell - \sum_{\ell,k} a_{\ell k}(q,t)\ \dot{q}_k\dot{q}_\ell + V(q,\ t) \tag{11-7}$$

$$H = T + V = E, \tag{11-8}$$

The Hamiltonian then equals the total energy.

If the constraints are time-varying, then the kinetic energy will contain terms arising from the velocities of the constraints. It will not be possible to represent the kinetic energy as a homogeneous quadratic form in this case, and $H \neq E$.

Finding the Hamiltonian - Examples

Consider the example of a simple mass-spring system in the absence of gravity.

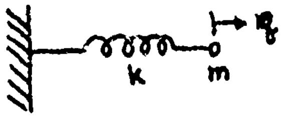


FIGURE 11-1

$$T = \frac{1}{2} m\dot{q}^2 \qquad V = \frac{k}{2} q^2 \qquad L = \frac{1}{2} m\dot{q}^2 - \frac{k}{2} q^2 \tag{11-9}$$

The momentum is

$$p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} \qquad \dot{q} = \frac{p}{m} \tag{11-10}$$

and hence

$$H = p\dot{q} - L = \frac{p^2}{m} - \frac{1}{2} m\dot{q}^2 + \frac{k}{2} q^2$$

$$H = \frac{p^2}{2m} + \frac{k}{2} q^2 \tag{11-11}$$

which is the desired Hamiltonian. The canonical equations are

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{m} \qquad \dot{p} = -\frac{\partial H}{\partial q} = -kq \tag{11-12}$$

As a second example consider a rotating pendulum. Let a bead of mass m be free to slide on a smooth circular wire of radius a. This wire hoop in turn rotates about a vertical axis with angular velocity ω. Gravity acts vertically downward. We desire to find H and derive the canonical equations.

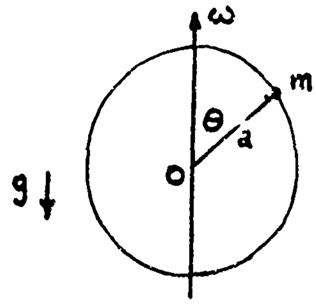


FIGURE 11-2

The kinetic and potential energies are

$$T = \frac{1}{2} m \left(a^2\dot{\theta}^2 + a^2\omega^2\sin^2\theta\right) \tag{11-13}$$

$$V = mga \cos\theta$$

$$L = T - V = \frac{1}{2} m \left(a^2\dot{\theta}^2 + a^2\omega^2\sin^2\theta\right) - mga \cos\theta\,.$$

The generalized momentum is

$$p = \frac{\partial L}{\partial \dot{\theta}} = ma^2\dot{\theta} \qquad \dot{\theta} = \frac{p}{ma^2}\,. \tag{11-14}$$

Hence

$$H = \frac{p^2}{2ma^2} - \frac{1}{2} ma^2\omega^2\sin^2\theta + mga\cos\theta \tag{11-15}$$

and the canonical equations become

$$\dot{\theta} = \frac{\partial H}{\partial p} = \frac{p}{ma^2} \tag{11-16}$$

$$\dot{p} = -\frac{\partial H}{\partial \theta} = ma^2\omega^2 \sin\theta\cos\theta + mga\sin\theta \tag{11-17}$$

Since t doesn't appear explicitly in L or H, H is a constant of the motion. Equation (11-13) shows that T is not in homogenous quadratic form so H is not the total energy. Examination of (11-15) verifies this.

In the Lagrange formalism a cyclic or "ignorable" coordinate is a coordinate q_i which is absent in the Lagrangian L. $\frac{\partial L}{\partial q_i} = 0$, and it follows from Lagrange's equation that the conjugate momentum is constant. If q_i does not appear in L, it will not appear in the Hamiltonian either. This is seen by noting that H and L differ only by $\sum_i p_i\dot{q}_i$ which does not contain q_i explicitly. Suppose in a problem with n generalized coordinates the ignorable coordinate is q_1, and the associated constant momentum is α_1. The Hamiltonian is then $H(q_2, q_3, \ldots q_n, \alpha_1, p_2, \ldots p_n)$.

The coordinate q_1 no longer appears in the problem, except as a constant in H, and in a real sense there are now only n-1 coordinates to be considered. In this respect the Hamiltonian formulation is superior to Lagrange's. In Lagrange's equations, ignorable coordinates do not dis-disappear so completely.

Because of the simplification of the Hamiltonian by ignorable coordinates, it is desirable to look for coordinate transformations which would make all the coordinates cyclic. If this could be done, the solution would be trivial. Thus given $H(q_i, p_i, t)$ where

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \qquad i = 1,2,\ldots n. \tag{11-18}$$

We desire to change from the 2n variables q_i, p_i to 2n new variables P_i, Q_i in such a way that it will be easy to determine the new equations of motion and further, we desire these new equtions be in canonical form also, i.e., we would like

$$\frac{dP_i}{dt} = -\frac{\partial H^1}{\partial Q_i} \qquad \frac{dQ_i}{dt} = \frac{\partial H^1}{\partial P_i} \tag{11-19}$$

where H^1 is easily related to H and is expressed as a function of P_i, Q_i and possibly t. According to a theorem of Jacobi, the desired transformation between the old (p_i, q_i) and new (P_i, Q_i) variables, which give rise to the new set of canonical equations (11-19) can be expressed as

$$p_i = \frac{\partial \overline{S}}{\partial q_i} \qquad P_i = -\frac{\partial \overline{S}}{\partial Q_i} \qquad \overline{S} = \overline{S}\,(q_i, Q_i, t) \tag{11-20}$$

This determining or generating function $\overline{S}$ must be expressed as a function of one set, either the q_i or the p_i, of the old variables, and

one set of the new. $\bar{S}$ must be chosen so that it is possible, by means of (11-20), to express the q_i, p_i in terms of the P_i, Q_i, or vica versa.

$$P_i = P_i\,(p_i,\ q_i,\ t) \tag{11-21}$$

$$Q_i = Q_i\,(p_i,\ q_i,\ t).$$

To do this let P_i and Q_i be continuous, with continuous first partial derivatives and

$$\frac{\partial\,(P_1,\ P_2,\ldots P_n,\ Q_1,\ Q_2,\ldots Q_n)}{\partial\,(p_1,\ p_2,\ldots p_n,\ q_1,\ q_2,\ldots q_n)} \not\equiv 0\,, \tag{11-22}$$

that is, the Jacobian must not vanish identically.

The expression implies the Jacobian determinant:

$$\begin{vmatrix}
\frac{\partial P_1}{\partial p_1} & \frac{\partial P_1}{\partial p_2} & \cdots & \frac{\partial P_1}{\partial p_n} & \frac{\partial P_1}{\partial q_1} & \cdots & \frac{\partial P_1}{\partial q_n} \\
\frac{\partial P_2}{\partial p_1} & \frac{\partial P_2}{\partial p_2} & \cdots & \frac{\partial P_2}{\partial p_n} & \frac{\partial P_2}{\partial q_1} & \cdots & \frac{\partial P_2}{\partial q_n} \\
\cdot & \cdot & & \cdot & \cdot & & \cdot \\
\cdot & \cdot & & \cdot & \cdot & & \cdot \\
\cdot & \cdot & & \cdot & \cdot & & \cdot \\
\frac{\partial P_n}{\partial p_1} & \cdots & \cdots & \frac{\partial P_n}{\partial p_n} & \frac{\partial P_n}{\partial q_1} & \cdots & \frac{\partial P_n}{\partial q_n} \\
\frac{\partial Q_1}{\partial p_1} & \cdots & \cdots & \frac{\partial Q_n}{\partial p_n} & \frac{\partial Q_n}{\partial q_1} & \cdots & \frac{\partial Q_n}{\partial q_n} \\
\vdots & & & \vdots & \vdots & & \vdots \\
\frac{\partial Q_n}{\partial p_1} & \cdots & \cdots & \frac{\partial Q_n}{\partial p_n} & \frac{\partial Q_n}{\partial q_1} & \cdots & \frac{\partial Q_n}{\partial q_n}
\end{vmatrix} \not\equiv 0$$

If (11-22) is satisfied, the functions P_i, Q_i are independent, i.e., there exists no function F such that

$$F(\overline{P}, \overline{Q}) = 0 \qquad (11\text{-}24)$$

This is a natural requirement. The new "coordinates" P_i, Q_i should be independent or the number of degrees of freedom would be reduced by the transformation.

In this transformation, the momenta p_i play much the same role as the coordinates q_i. This is one of the underlying ideas of the Hamiltonian methods, and is closely allied with "State space" methods in modern differential equations theory and automatic control.

If such a transformation is applied to the canonical equations, in general, one obtains

$$\dot{P}_i = G_1(\overline{P}, \overline{Q}, t); \; \dot{Q}_i = G_2(\overline{P}, \overline{Q}, t) \quad i = 1, 2, \ldots n$$

where G_1 and G_2 are two functions whose nature depends on the transformation used.

If, however, the transformation is of such nature that the new equations have the same <u>form</u> as the canonical equations, i.e.,

$$P_i = -\frac{\partial H^1}{\partial Q_i} \qquad \dot{Q}_i = \frac{\partial H^1}{\partial P_i} \quad i = 1, 2, \ldots n$$

for some $H^1 = H^1(\overline{P}, \overline{Q}, t)$, then it is called a <u>canonical transformation</u>.

But to return to the theorem of Jacobi, this desired relationship between the old and new variables gives a new Hamiltonian, H^1, as

$$H^1 = H + \frac{\partial S}{\partial t} = H^1(P_1, Q_1, t) \tag{11-25}$$

so long as (11-20) is true. We can thus easily determine the new force equations from a transformation of variables. Now let's prove the Jacobian transformation theorem.

12. JACOBIAN TRANSFORMATION THEOREM

Given

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i} \quad ; \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad ; \qquad i = 1, 2, \ldots n, \tag{12-1}$$

these have the solutions

$$\begin{aligned} q_i &= q_i\,(t,\ a_1,\ a_2,\ \ldots\ a_{2n}) \\ p_i &= p_i\,(t,\ a_1,\ a_2,\ \ldots\ a_{2n}) \end{aligned} \tag{12-2}$$

and

$$\frac{dQ_i}{dt} = \frac{\partial H^1}{\partial P_i} \qquad \frac{dP_i}{dt} = -\frac{\partial H^1}{\partial Q_i} \tag{12-3}$$

have the solutions

$$\begin{aligned} Q_i &= Q_i\,(t,\ a_1,\ a_2,\ \ldots\ a_{2n}) \\ P_i &= P_i\,(t,\ a_1,\ a_2,\ \ldots\ a_{2n}) \end{aligned} \tag{12-4}$$

where $a_1, a_2, \ldots a_{2n}$ are the arbitrary constants of the solutions.

To proceed we denote certain variations or operations. The symbol d attached to any function, such as dq_i, will denote that when the function has been expressed in terms of t and a_r, it is t alone which is varied;

while the symbol δ, such as δq, implies that any or all of the a_r are varied but that t is not changed. These are virtual displacements. We thus have the following:

$$dq_i = \frac{\partial q_i}{\partial t} dt \qquad \delta q_i = \sum_r \frac{\partial q_i}{\partial a_r} da_r \quad r = 1, 2, \cdots 2n. \tag{12-5}$$

For a function $\bar{S} = \bar{S}\,(q_i, Q_i, t)$ one thus has

$$d\bar{S} = \sum \left(\frac{\partial \bar{S}}{\partial q_i} dq_i + \frac{\partial \bar{S}}{\partial Q_i} dQ_i \right) + \frac{\partial \bar{S}}{\partial t} dt \tag{12-6}$$

as the definition of the total differential and

$$\delta \bar{S} = \sum \left(\frac{\partial \bar{S}}{\partial q_i} \delta q_i + \frac{\partial \bar{S}}{\partial Q_i} \delta Q_i \right) \tag{12-7}$$

for the virtual displacement of $\bar{S}$.

Since the variations denoted by d and δ are independent, the commutative law holds, i.e., $\delta \cdot d$ is the same as $d \cdot \delta$ operation.

With this introduction, the proof that equation (11-20) of the last section transforms (12-1) into (12-3) then follows:

First multiply $\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}$ and $\frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}$ by δp_i and $-\delta q_i$ respectively and add for all values of i, to obtain

$$\sum \left(\frac{dq_i}{dt} \delta p_i - \frac{dp_i}{dt} \delta q_i \right) = \sum \left(\frac{\partial H}{\partial p_i} \delta p_i + \frac{\partial H}{\partial q_i} \delta q_i \right) = \delta H. \tag{12-8}$$

A similar process applied to $p_i = \frac{\partial \bar{S}}{\partial q_i}$ and $P_i = -\frac{\partial \bar{S}}{\partial Q_i}$ gives

$$\sum \left(p_i \, \delta q_i - P_i \, \delta Q_i \right) = \sum \left(\frac{\partial \bar{S}}{\partial q_i} \delta q_i + \frac{\partial \bar{S}}{\partial Q_i} \delta Q_i \right) = \delta \bar{S} \tag{12-9}$$

Using the d operation instead of the δ operation, the last equation becomes

$$\sum \left(p_i dq_i - P_i dQ_i \right) = \sum \left(\frac{\partial \bar{S}}{\partial q_i} dq_i + \frac{\partial \bar{S}}{\partial Q_i} dQ_i \right) = d\bar{S} - \frac{\partial \bar{S}}{\partial t} dt \tag{12-10}$$

Since $\bar{S}$ may contain t explicitly as well as through q_i and Q_i we must include $\frac{\partial \bar{S}}{\partial t}$ in the total differential for $\bar{S}$. This equation (12-10) is sometimes used to define a canonical or contact transformation. We shall refer to it again.

Now operate on equation (12-9) with the operator d and on equation (12-10) with the operator δ and subtract the two results. Since we have $d\delta q_i = \delta dq_i$ and $d\delta\bar{S} = \delta d\bar{S}$, all terms in which both δ and d operate on the same function will cancel when we subtract the two results. From (12-9) we have $d(p_i \, \delta q_i) - d(P_i \, \delta Q_i) = d\delta\bar{S}$ or

$$dp_i \delta q_i + p_i d\delta q_i - dP_i \delta Q_i - P_i d\delta Q_i = d\delta\bar{S} \tag{12-11}$$

and from (12-10) we obtain

$$\delta(p_i dq_i) - \delta(P_i dQ_i) = \delta d\bar{S} - \delta\left(\frac{\partial \bar{S}}{\partial t} dt\right)$$

$$\delta p_i dq_i + p_i \delta dq_i - \delta P_i dQ_i - P_i \delta dQ_i = \delta d\bar{S} - \delta \frac{\partial \bar{S}}{\partial t} dt - \frac{\partial \bar{S}}{\partial t} \delta dt \tag{12-12}$$

but $\delta dt = 0$ by definition and then subtracting (12-12) from (12-11) gives

$$\Sigma(dp_i\delta q_i - dq_i\delta p_i) - \Sigma(dP_i\delta Q_i - dQ_i\delta P_i) = dt\delta\left[\frac{\partial\bar{S}}{\partial t}\right] \tag{12-13}$$

which can be written

$$\sum\left(\frac{dp_i}{dt}\delta q_i - \frac{dq_i}{dt}\delta p_i\right) + \sum\left(\frac{dQ_i}{dt}\delta P_i - \frac{dP_i}{dt}\delta Q_i\right) = \delta\left(\frac{\partial\bar{S}}{\partial t}\right) \tag{12-14}$$

Recall equation (12-8)

$$\sum\left(\frac{dq_i}{dt}\delta p_i - \frac{dp_i}{dt}\delta q_i\right) = \delta H . \tag{12-8}$$

Upon adding these last two equations we have

$$\sum\left(\frac{dQ_i}{dt}\delta P_i - \frac{dP_i}{dt}\delta Q_i\right) = \delta\left(\frac{\partial\bar{S}}{\partial t} + H\right) = \delta H^1 \tag{12-15}$$

where we define

$$H^1 = H + \frac{\partial\bar{S}}{\partial t} = H^1\ (P_i,\ Q_i,\ t)\,. \tag{12-16}$$

From this we can write

$$\delta H^1 = \sum\left(\frac{\partial H^1}{\partial P_i}\delta P_i + \frac{\partial H^1}{\partial Q_i}\delta Q_i\right) \tag{12-17}$$

and then equation (12-15) becomes,

$$\sum\left(\frac{dQ_i}{dt}\delta P_i - \frac{dP_i}{dt}\delta Q_i\right) = \sum\left(\frac{\partial H^1}{\partial P_i}\delta P_i + \frac{\partial H^1}{\partial Q_i}\delta Q_i\right)$$

or

$$\left(\frac{dQ_i}{dt} - \frac{\partial H^1}{\partial P_i}\right)\delta P_i = \left(\frac{dP_i}{dt} + \frac{\partial H^1}{\partial Q_i}\right)\delta Q_i \qquad (12\text{-}18)$$

Since δP_i and δQ_i are all linearly independent variables, it follows that

$$\frac{dQ_i}{dt} = \frac{\partial H^1}{\partial P_i} \quad \text{and} \quad \frac{dP_i}{dt} = -\frac{\partial H^1}{\partial Q_i} \qquad \text{Q.E.D.} \qquad (12\text{-}19)$$

This transformation theorem provides us with a means of changing from one set of variables to another by means of a generating function $\bar{S}$. To summarize, if there exists a function $\bar{S}$, with continuous first partial derivatives such that for $\bar{S} = \bar{S}(q_i, Q_i, t)$,

$$p_i = \frac{\partial \bar{S}}{\partial q_i} \quad , \quad P_i = -\frac{\partial \bar{S}}{\partial Q_i} \qquad (12\text{-}20)$$

which implies that

$$\Sigma P_i dQ_i - \sum_i p_i dq_i = -d\bar{S} + \frac{\partial \bar{S}}{\partial t} dt \qquad (12\text{-}21)$$

(see equation 12-10), then the transformed equations are

$$\frac{dQ_i}{dt} = \frac{\partial H^1}{\partial P_i} \qquad \frac{dP_i}{dt} = -\frac{\partial H^1}{\partial Q_i} \qquad (12\text{-}22)$$

$$H^1 = H + \frac{\partial \bar{S}}{\partial t}\ .$$

and the transformation is canonical.

The equations of transformation depend upon the choice of $\overline{S}$ and as was mentioned, $\overline{S}$ can be any function of one of the old and one of the new variables. Particular forms of $\overline{S}$ are of interest. For example if we choose $\overline{S} = S$ such that $\frac{\partial H^1}{\partial P_i} = \frac{\partial H^1}{\partial Q_i} = 0$, then (12-22) gives $\frac{dP_i}{dt} = 0$, $\frac{dQ_i}{dt} = 0$ and hence

$$P_i = \text{constants} \qquad\qquad Q_i = \text{constants}$$

and the problem would be solved by a mere transformation. All coordinates would become ignorable ones. Before we develop techniques for obtaining S, let us consider some examples of transformations.

Consider the function

$$\overline{S} = \sum_i q_i Q_i \ . \tag{12-23}$$

Note that

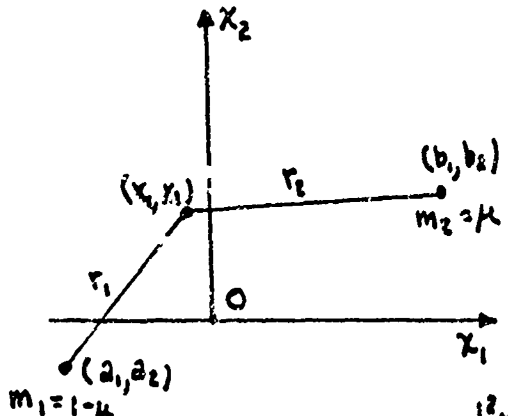
$$\frac{\partial \overline{S}}{\partial q_i} = Q_i ; \ \frac{\partial \overline{S}}{\partial Q_i} = q_i \ .$$

Combining these with equation (12-20), we have

$$p_i = Q_i ; \ - P_i = q_i \tag{12-24}$$

The canonical transformation which is generated by this $\overline{S}$ interchanges the role of momenta and coordinates, emphasizing the equal standing of those two types of quantities in Hamiltonian theory.

Now consider another example, the planar restricted three-body problem.



$$V = -\left(\frac{1-\mu}{r_1} + \frac{\mu}{r_2}\right)$$

$$\frac{d^2 x_i}{dt^2} = \frac{(1-\mu)(a_i - x_i)}{r_1^3} + \frac{\mu(b_i - x_i)}{r_2^3} = -\frac{\partial V}{\partial x_i}$$

FIGURE 12-1

The equations of motion are

$$\frac{dx_i}{dt} = y_i = \frac{\partial H}{\partial y_i} \quad ; \quad \frac{dy_i}{dt} = -\frac{\partial V}{\partial x_i} = -\frac{\partial H}{\partial x_i}$$

$$H = \frac{1}{2}(y_1^2 + y_2^2) + V \quad ; \quad y_1 = \dot{x}_1 \quad ; \quad y_2 = \dot{x}_2$$

We are going to change from old coordinates (x_i, y_i) to a new set (q_i, p_i), the new set will be canonical and (incidentally) will be those of a rotating coordinate system. We select $\overline{S} = \overline{S}(x_i, p_i, t)$ as

$$\overline{S} = x_1 (p_1 \cos t - p_2 \sin t) + x_2 (p_1 \sin t + p_2 \cos t)$$

from which we form

$$y_1 = \frac{\partial \bar{S}}{\partial x_1} = p_1 \cos t - p_2 \sin t; \qquad y_2 = \frac{\partial \bar{S}}{\partial x_2} = p_1 \sin t + p_2 \cos t$$

$$q_1 = \frac{\partial \bar{S}}{\partial p_1} = x_1 \cos t + x_2 \sin t; \qquad q_2 = \frac{\partial \bar{S}}{\partial p_2} = - x_1 \sin t + x_2 \cos t \ .$$

The new Hamiltonian becomes

$$H^1 = H + \frac{\partial \bar{S}}{\partial t} = \frac{1}{2} (y_1^2 + y_2^2) + V + x_1 (-p_1 \sin t - p_2 \cos t)$$
$$+ x_2 (p_1 \cos t - p_2 \sin t) \ .$$

To eliminate x_1 and x_2 we form

$$q_1 \cos t = x_1 \cos^2 t + x_2 \cos t \sin t$$

$$-q_2 \sin t = x_1 \sin^2 t - x_2 \cos t \sin t$$

then adding and subtracting we obtain

$$x_1 = q_1 \cos t - q_2 \sin t$$

(see page 92)

$$x_2 = q_1 \sin t + q_2 \cos t \ .$$

Note that this is the transformation from fixed (x_1, x_2) to rotating coordinate systems (q_1, q_2).

If we substitute this into H^1 we obtain

$$H^1 = \frac{1}{2}(p_1^2 + p_2^2) + V + p_1q_2 - q_1p_2$$

and the equations of motion are

$$\frac{dq_1}{dt} = \frac{\partial H^1}{\partial p_1} = p_1 + q_2 \qquad \frac{dq_2}{dt} = \frac{\partial H^1}{\partial p_2} = p_2 - q_1$$

$$\frac{dp_1}{dt} = -\frac{\partial H^1}{\partial q_1} = -\frac{\partial V}{\partial q_1} + p_2 \qquad \frac{dp_2}{dt} = -\frac{\partial H^1}{\partial q_2} = -\frac{\partial V}{\partial q_1} - p_1$$

Note that

$$\frac{d^2q_1}{dt^2} = \frac{dp_1}{dt} + \frac{dq_2}{dt} = -\frac{\partial V}{\partial q_1} + 2p_2 - q_1$$

$$\frac{d^2q_2}{dt^2} = \frac{dp_2}{dt} - \frac{dq_1}{dt} = -\frac{\partial V}{\partial q_2} - 2p_1 - q_2$$

which checks previous results of equation (7-4).

The procedure here was to pick an $\bar{S}$ function $\bar{S} = \bar{S}(x_i, p_i, t)$ and then form the new Hamiltonian H^1 by

$$H^1 = H + \frac{\partial \bar{S}}{\partial t} = H^1(x_i, y_i, p_i, t)$$

we then used

$$q_i = \frac{\partial \bar{S}}{\partial p_i}; \qquad y_i = \frac{\partial \bar{S}}{\partial x_i}$$

from equation (12-39), which we shall shortly prove, in order to eliminate x_i and y_i to form

$$H^1 = H^1 (q_i, p_i, t) .$$

The new canonical equations of motion are then formed directly as

$$\frac{dq_i}{dt} = \frac{\partial H^1}{\partial p_i}; \qquad \frac{dp_i}{dt} = -\frac{\partial H^1}{\partial q_i} .$$

Now consider the simple mass-spring system of equation (11-9), the harmonic oscillator. For this case we choose

$$\overline{S} = \overline{S} (q, Q) = \frac{\sqrt{mk}}{2} q^2 \cot Q \quad . \tag{12-25}$$

This is a generating function for a one-degree-of-freedom system which has a single generalized coordinate q. Applying equations (12-20),

$$\begin{aligned} p &= \frac{\partial \overline{S}}{\partial q} = \sqrt{mk}\, q \cot Q \\ P &= -\frac{\partial \overline{S}}{\partial Q} = \frac{\sqrt{mk}}{2} \frac{q^2}{\sin^2 Q} \quad . \end{aligned} \tag{12-26}$$

These equations may be solved for the old coordinates in terms of the new

$$\begin{aligned} p &= \sqrt{2\sqrt{mk}\, P} \cos Q \\ q &= \sqrt{\frac{2P}{\sqrt{mk}}} \sin Q \end{aligned} \tag{12-27}$$

It is also possible to solve for Q, P in terms of p, q.

$$P = \frac{p^2}{2\sqrt{mk}} + \frac{q^2\sqrt{mk}}{2}$$

$$Q = \tan^{-1}\left\{\frac{pq\sqrt{mk}}{p^2 + mkq^2}\right\} . \tag{12-28}$$

Now consider the harmonic oscillator of equation (11-11). The Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{k}{2} q^2 . \tag{12-29}$$

We can express this Hamiltonian in terms of the new coordinates, by substituting (12-27) into (12-29). The result is

$$H = P\sqrt{\frac{k}{m}} \tag{12-30}$$

Since $\overline{S}$ does not contain t, $\frac{\partial \overline{S}}{\partial t} = 0$, and, from (12-22)

$$H^1 = H = P\sqrt{\frac{k}{m}} . \tag{12-31}$$

Note that the transformed Hamiltonian is cyclic in Q. The transformed canonical equations are, by (12-22):

$$\dot{Q} = \frac{\partial H}{\partial P} = \sqrt{\frac{k}{m}} \qquad \dot{P} = -\frac{\partial H}{\partial Q} = 0 , \tag{12-32}$$

which may readily be integrated to give

$$Q = \alpha + \sqrt{\frac{k}{m}}\, t;\quad P = \beta = \text{constant.} \tag{12-33}$$

The motion of the original problem is obtained by substituting these expressions into (12-27).

$$\begin{aligned} p &= \sqrt{2\sqrt{mk}\,\beta}\,\cos\left\{\sqrt{\frac{k}{m}}\, t + \alpha\right\} \\ q &= \sqrt{\frac{2\beta}{\sqrt{mk}}}\,\sin\left\{\sqrt{\frac{k}{m}}\, t + \alpha\right\}\ . \end{aligned} \tag{12-34}$$

The α, β constants are determined from initial conditions. This fortuitous choice of a generating function $\bar{S}$ has transformed the problem to one which could easily integrated. It would be nice to have a systematic way to pick such an $\bar{S}$ function. Before discussing that point, let us consider other $\bar{S}$ function in more detail.

In making the last transformation we used as a generating function the quantity $\bar{S} = \bar{S}\ (q_i, Q_i, t)$. Actually we are free to choose other variables, one from each side of the transformation. For example we could find the canonical equation just as well by choosing $\bar{S}_2 = \bar{S}_2(q_i, P_i, t)$ (which is what was used in the restricted three body example). The proof of the Hamilton-Jacobi Transformation theorem would proceed precisely as before (mutatis mutandis) with this new generating function $\bar{S}_1$.

The best way to change to other variables is to make judicious use of equation (12-10). The canonical transformation was generated by $\bar{S}_1 = \bar{S}_1\ (q_i, Q_i, t)$ and therefore equation (12-10) gives

$$\sum_i P_i dQ_i - \sum_i p_i dq_i - \frac{\partial \bar{S}_1}{\partial t} dt = - d\bar{S}_1 \tag{12-35}$$

and

$$p_i = \frac{\partial \bar{S}_1}{\partial q_i} ; \qquad P_i = - \frac{\partial \bar{S}_1}{\partial Q_i} . \tag{12-36}$$

Now define a new function

$$\bar{S}_2 \ (q_i, P_i, t) = \bar{S}_1 \ (q_i, Q_i, t) + \sum_i P_i Q_i . \tag{12-37}$$

Using this definition,

$$\begin{aligned} d\bar{S}_1 &= d\bar{S}_2 - \sum_i P_i dQ_i - \sum_i Q_i dP_i \\ \frac{\partial \bar{S}_1}{\partial t} &= \frac{\partial \bar{S}_2}{\partial t} . \end{aligned} \tag{12-38}$$

Substituting these relations into (12-35), and simplifying

$$- \sum_i Q_i dP_i - \sum_i p_i dq_i - \frac{\partial \bar{S}_2}{\partial t} dt = - d\bar{S}_2 = - \frac{\partial \bar{S}_2}{\partial P_i} dP_i - \frac{\partial \bar{S}_2}{\partial q_i} dq_i - \frac{\partial S_2}{\partial t} dt ,$$

so that

$$Q_i = \frac{\partial \bar{S}_2}{\partial P_i} ; \qquad p_i = \frac{\partial \bar{S}_2}{\partial q_i} \tag{12-39}$$

By means of these relations, a canonical transformation may be generated just as before.

Goldstein (page 240 et. seq.) considers additional change of variables such as $\overline{S}_3$ (p_i, Q_i, t) and $\overline{S}_4$ (p_i, P_i, t).

Now let us return to a search for a formal procedure for finding a special generating function, S, so that the new "variables" are all constants.

13. HAMILTON-JACOBI PARTIAL DIFFERENTIAL EQUATION

Now let's consider the problem of picking a particular generating function so all of the coordinates are ignorable. First note that if the new Hamiltonian, H^1, is a constant or at least not a function of Q_i or P_i, then

$$\frac{dP_i}{dt} = -\frac{\partial H^1}{\partial Q_i} = 0 ; \qquad \frac{dQ_i}{dt} = \frac{\partial H^1}{\partial P_i} = 0 \qquad i = 1,2,\ldots n \qquad (13\text{-}1)$$

and integration of these equations gives

$$P_i = \beta_i; \qquad Q_i = \alpha; \qquad i = 1,2,\ldots n. \qquad (13\text{-}2)$$

Using the transformation equations, like (12-39), we can evaluate p_i (P_i, Q_i, t) and q_i (P_i, Q_i, t) and the problem is solved.

If the new Hamiltonian is to be a constant, say zero, then from equation (12-22) we require

$$H^1 = \frac{\partial S}{\partial t} + H(q_i, p_i, t) = 0 \qquad (13\text{-}3)$$

This equation holds true whether we use $S = S_1$ (q_i, Q_i, t) or $S = S_2$ (q_i, P_i, t) and in either case

$$p_i = \frac{\partial S}{\partial q_i} \qquad (13\text{-}4)$$

Substituting this into (13-3),

$$\frac{\partial S}{\partial t} + H\left(q_1, q_2, \ldots q_n, \frac{\partial S}{\partial q_1}, \frac{\partial S}{\partial q_2}, \ldots \frac{\partial S}{\partial q_n} t\right) = 0. \tag{13-5}$$

This is a partial differential equation for the function S. It is called the Hamilton-Jacobi partial differential equation (PDE) and is a first order nonlinear differential equation in the n + 1 independent variables $q_1, q_2, \ldots q_n, t$.

Since there are n + 1 independent variables, there would arise in the solution of the equation n + 1 arbitrary constants of integration. One of these will, however, be additive. Since only derivatives of S appear in the equation, then if S is a solution of the equation, $S + \alpha_o$ is also a solution. There must then be n non-additive constants of integration. The solution is

$$S = S\,(q_1, q_2, \ldots q_n, \alpha_1, \alpha_2, \ldots \alpha_n, t)$$

where the dependence on the n non additive constants, α_n, must be non degenerative. We expect to deal with functions like $S_1\,(q_i, Q_i, t)$ or $S_2\,(q_i, P_i, t)$. However, note that if the Hamilton-Jacobi equation is satisfied, $H^1 = 0$ and both P_i and Q_i are themselves arbitrary constants of integration. Accordingly, once we obtain a solution of the form $S\,(q_i, \alpha_i, t)$, $i = 1,2, \ldots n$, we may associate the α_i's with either the P_i or Q_i and write

$$S_1 = S\,(q_1 \cdots q_n,\ Q_1 \cdots Q_n,\ t)$$

or (13-6)

$$S_2 = S\,(q_1 \cdots q_n,\ P_1 \cdots P_n,\ t).$$

These two functions will, of course, generate different transformations, but either may be chosen as desired.

This is illustrative of the considerable flexibility that one has in picking arbitrary constants in Hamilton-Jacobi problems. Almost any unambiguous choice may be made.

To summarize this method we perform the following steps:

1. Determine the Hamiltonian of the physical system (see page 106),

$$H\,(q_i,\ p_i,\ t) \quad \text{where} \quad \dot{q}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

2. Form the Hamilton-Jacobi partial differential equation.

$$\frac{\partial S}{\partial t} + H = 0$$

3. Solve the Hamilton-Jacobi equation, and the n non-additive constants of integration will be introduced. The solution is

$$S\,(q_1 \cdots q_n,\ \alpha_1,,\alpha_n,\ t).$$

4. Identify the constants of integration with either the P_i or Q_i.

5. If the constants are identified with Q_i, then derive the transformation equations from:

$$\frac{\partial S_1}{\partial q_i} = p_i \; ; \qquad \frac{\partial S_1}{\partial Q_i} = - P_i \quad ; \quad S_1 \,(q_i, Q_i, t) \,.$$

If the constants are identified with the P_i, derive the transformation relations from:

$$\frac{\partial S_2}{\partial q_i} = p_i \quad ; \quad \frac{\partial S_2}{\partial P_i} = Q_i \quad ; \quad S_2 \,(q_i, P_i, t)$$

6. Introduce n more arbitrary constants for the values of the P_i or Q_i, whichever were not determined in step 5.

7. With P_i and Q_i completely determined by the 2n constants of integration, use the transformation equations to obtain the solution to the original problem

$$q_i = q_i \,(\overline{P}, \overline{Q}, t) = q_i \,(\alpha_1 \cdots \alpha_n, \beta_1 \cdots \beta_n, t); i = 1, 2, \cdots n$$

$$p_i = p_i \,(\overline{P}, \overline{Q}, t) = p_i \,(\alpha_1 \cdots \alpha_n, \beta_1 \cdots \beta_n, t); i = 1, 2. \cdots n \,.$$

To illustrate this procedure, consider the example of the harmonic oscillator. The Hamiltonian was derived earlier as (see Equation 11-11),

$$H = \frac{p^2}{2m} + \frac{k}{2} q^2 \qquad (13\text{-}7)$$

Now from the Hamilton-Jacobi equation, replacing p by $\frac{\partial S}{\partial q}$ in the Hamiltonian

$$\frac{\partial S}{\partial t} + \frac{1}{2m} \left(\frac{\partial S}{\partial q}\right)^2 + \frac{k}{2} q^2 = 0 \ .$$

To solve this equation, we assume it is separable, i.e.

$$S = W\ (q) + T\ (t)$$

where W is a function of q alone, T is a function of t alone. Substituting this expression into the original equation (13-8):

$$\frac{kq^2}{2} + \frac{1}{2m} \left(\frac{\partial W}{\partial q}\right)^2 = - \frac{\partial T}{\partial t} \ . \tag{13-9}$$

The left side is a function of q alone, the right is a function of t alone. (Note the partials in this case are actually ordinary derivatives.) Thus both sides of (13-9) must equal some (positive) constant.

$$\frac{kq^2}{2} + \frac{1}{2m} \left(\frac{dW}{dq}\right)^2 = \frac{k}{2} \alpha^2 \tag{13-10}$$

$$- \frac{dT}{dt} = \frac{k}{2} \alpha^2 \ . \tag{13-11}$$

The form of the constant of integration has been chosen somewhat arbitrarily, to simplify subsequent calculations. From equation (13-11)

$$\frac{dT}{dt} = -\frac{k}{2}\alpha^2 . \tag{13-12}$$

Integrating this,

$$T = -\frac{k}{2}\alpha^2 t + \alpha_2 , \tag{13-13}$$

α_2 is the additive constant, which may be set equal to zero without any loss. From equation (13-10)

$$\frac{dW}{dq} = \sqrt{mk}\ \ \sqrt{\alpha^2 - q^2} \tag{13-14}$$

integrating,

$$W = \sqrt{mk}\ \int_o^q \sqrt{\alpha^2 - x^2}\, dx + \alpha_3 \tag{13-15}$$

The additive constant α_3 may be dropped. Carrying out the integration,

$$W = \frac{\sqrt{mk}}{2}\,[q\sqrt{\alpha^2 - q^2} + \alpha^2 \sin^{-1}\frac{q}{\alpha}] , \tag{13-16}$$

Thus the solution of the Hamilton-Jacobi PDE is

$$S = W + T = \frac{\sqrt{mk}}{2}\,[q\ \sqrt{\alpha^2 - q^2} + \alpha^2 \sin^{-1}\frac{q}{\alpha}] - \frac{k}{2}\alpha^2 t , \tag{13-17}$$

This satisfies the equation, and contains the one expected non-additive constant of integration α.

In this case we choose to identify the constant α with the new (constant) coordinate Q, so that

$$S_1\ (q,\ Q,\ t) = \frac{\sqrt{mk}}{2}\ [q\sqrt{Q^2 - q^2} + Q^2 \sin^{-1}\frac{q}{Q}] - \frac{k}{2}\,Q^2 t\ . \tag{13-18}$$

The transformation equations are derived from:

$$p = \frac{\partial S_1}{\partial q} = \sqrt{mk}\ \sqrt{Q^2 - q^2} \tag{13-19}$$

$$- P = \frac{\partial S_1}{\partial Q} = \sqrt{mk}\ \ Q \sin^{-1}\frac{q}{Q} - kQt \tag{13-20}$$

Solving these equations for p, q, in terms of P, Q, t

$$q = Q \sin\left\{\sqrt{\frac{k}{m}}\,t - \frac{P}{Q\sqrt{mk}}\right\} \tag{13-21}$$

$$p = \sqrt{mk}\ Q \cos\left\{\sqrt{\frac{k}{m}}\,t - \frac{P}{Q\sqrt{mk}}\right\} \tag{13-22}$$

Since P and Q are arbitrary constants of integration and so may be determined by initial conditions, the problem is solved.

To verify that this transformation does what is expected, we calculate the new Hamiltonian H^1 (P, Q, t).

$$H^1 = H + \frac{\partial S}{\partial t} = \frac{p^2}{2m} + \frac{k}{2}q^2 - \frac{k}{2}\,Q^2\ ,$$

Substituting for p and q from (13-21) and (13-22) we obtain

$$H^1 = 0$$

as expected, and thus $\dot{P} = \dot{Q} = 0$, and P and Q are indeed constant. To emphasis that these are constants we could consider a transformation for q_i, p_i to an α_i, β_i coordinate system

14. SEPARABILITY OF HAMILTON-JACOBI EQUATION

To carry out this procedure we need to solve a partial differential equation. We have traded the solution of ordinary differential equations for a partial differential equation. This might be thought to be a hinderance rather than a help in solving physical problems. The solution of partial differential equations is a sometimes thing.

However, many problems in physics, and especially in celestial mechanics lead to Hamilton-Jacobi equations which are separable. When the equation has this property, it may be easily solved.

Suppose we assume a solution to the Hamilton-Jacobi equation which is a sum of functions of a single variable:

$$S(q_1 .. q_n, t) = S_1(q_1) + S_2(q_2) + \dots S_n(q_n) + S_t(t) \,. \tag{14-1}$$

If, on substituting this solution back into the Hamilton-Jacobi equation, it may be factored into separate functions

$$f_1\left(\frac{dS_1}{dq_1}, q_1\right) + f_2\left(\frac{dS_2}{dq_2}, q_2\right) + \dots f_n\left(\frac{dS_n}{dq_n}, q_n\right) + f_t\left(\frac{dS_T}{dt}, t\right) = 0 \tag{14-2}$$

then the equation is said to be <u>separable</u>. If the equation can be factored in this way, each of the functions must be equal to a constant.

$$f_i\left(\frac{dS_i}{dq_i}, q_i\right) = \alpha_i \quad i = 1, 2, \dots n \tag{14-3}$$

$$f_t\left(\frac{dS_t}{dt}, t\right) = \alpha_t \tag{14-4}$$

but not all these α's are independent, for

$$\sum_{i=1}^{n} \alpha_i + \alpha_t = 0 \tag{14-5}$$

in order for the equation to be satisfied. The partial differential equation in n+1 independent variables is thus reduced to n+1 first-order <u>ordinary</u> differential equations, each of which may always be reduced to quadratures.

It is of interest to determine what is the most general functional form of H for which the system will be separable. An excellent discussion of this is found in the book by Pars on page 291 and again on page 320.

Liouville's system is a rather general form which represents a separable orthogonal system. If the kinetic energy can be represented by

$$T = \frac{1}{2}\left(X_1 + X_2 + \cdots X_n\right)\left[\frac{\dot{q}_1^2}{R_1} + \frac{\dot{q}_2^2}{R_2} + \cdots \frac{\dot{q}_n^2}{R_n}\right] \tag{14-6}$$

or equivalently,

$$T = \frac{1}{2(X_1 + X_2 + \cdots X_n)}\left[R_1 p_1^2 + R_2 p_2^2 + \cdots R_n p_n^2\right] \tag{14-7}$$

and

$$V = \frac{\xi_1 + \xi_2 + \cdots \xi_n}{X_1 + X_2 + \cdots X_n}\,, \tag{14-8}$$

where X_n, R_n and ξ_n are functions of q_n, then the Hamilton-Jacobi PDE will be separable.

This Liouville's system is not, however, the most general separable orthogonal system. The most general orthogonal system is given by Stäckel's theorem (Pars page 321) which we shall not approach here.

It should be clearly noted that separability is a property jointly of the system <u>and</u> the chosen coordinates. Some physical systems have a Hamiltonian which leads to a separable Hamilton-Jacobi PDE if rectangular coordinates are used but will not separate if written in spherical coordinates. For example the Stark effect where $V = -\frac{\mu}{r} + gx$ is separable in parabolic coordinates where $u = \frac{1}{2}(r+x)$, $v = \frac{1}{2}(r-x)$. The two fixed center problem separates in confocal coordinates.

This separability condition has been used by Vinti, Garfinkel, Sterne and Barrar to great advantage in solving the oblate earth problem. See Deutsch's book page 185, et. seq.

should be clear that the times when one is able to solve the Hamilton-Jacobi PDE are rather rare. The solutions to partial differential equations are a sometimes thing. We can, however, consider the potential function to be broken into two parts, one due to a single body and the rest as a perturbation due to the other planets

$$V = V_o + \hat{R}. \tag{15-1}$$

The Hamiltonian can then be written

$$H = T + V = T + V_o + \hat{R} = H_o + \hat{R} \tag{15-2}$$

$$H_o = T + V_o \, . \tag{15-3}$$

The canonical equations are then

$$\frac{dq_i}{dt} = \frac{\partial(H_o + \hat{R})}{\partial p_i} \quad ; \quad \frac{dp_i}{dt} = -\frac{\partial(H_o + \hat{R})}{\partial q_i} \tag{15-4}$$

Now let's transform to a new set of canonical variables, which we choose to call α_i, β_i, by means of

$$p_i = \frac{\partial S}{\partial q_i} \qquad \beta_i = \frac{\partial S}{\partial \alpha_i} \tag{15-5}$$

where $S = S(q_i, \alpha_i, t)$ is the solution to the PDE,

$$H_o\left(q_i, \frac{\partial S}{\partial q_i}, t\right) + \frac{\partial S}{\partial t} = 0 \; . \tag{15-6}$$

By the Jacobi transformation theorem we know we will obtain a new canonical set

$$\frac{d\alpha_i}{dt} = -\frac{\partial H^1}{\partial \beta_i} \; ; \; \frac{d\beta_i}{dt} = \frac{\partial H^1}{\partial \alpha_i} \tag{15-7}$$

with

$$H^1 = H_o + \hat{R} + \frac{\partial S}{\partial t} = \hat{R} \tag{15-8}$$

since

$$H_o + \frac{\partial S}{\partial t} = 0 \; . \tag{15-9}$$

Hence

$$\frac{d\alpha_i}{dt} = -\frac{\partial \hat{R}}{\partial \beta_i} \quad \frac{d\beta_i}{dt} = \frac{\partial \hat{R}}{\partial \alpha_i} \; , \tag{15-10}$$

Thus we solve the equations with $\hat{R} = 0$, which amounts to solving equation (15-6) for S, and obtain a canonical set of <u>constants</u> α_i, β_i. These are "constants" when $\hat{R} = 0$. These "constants" are then considered as variables in equation (15-10) so we speak of this as the variation of constants or variation of parameters method.

Actually the equations with $\hat{R} = 0$ are solved mainly to indicate the choice of the new variables and are not used unless we know in advance that a solution can be obtained.

As an example let's apply this technique to a planet in the solar system. The H_0 will be that due to the sun and planet alone, while the other planets will give rise to the perturbations. We use the variables α_i, β_i, instead of P_i, Q_i, as before, in order to emphasize the fact that these new variables, α_i, β_i are really constants in the H_0 solution.

For n-bodies we showed that (see page 63),

$$\frac{d^2\bar{r}_i}{dt^2} + \frac{k^2(m_i+m_0)}{r_i^3}\bar{r}_i = \text{grad}\,\hat{R}_i \tag{15-11}$$

$$\hat{R}_i = k^2 \sum_{j=1}^{n-1} m_j \left(\frac{1}{r_{ij}} - \frac{\bar{r}_i \cdot \bar{r}_j}{r_j^3} \right). \tag{15-12}$$

$\hat{R}_i$ is called the disturbing function. Consider a typical case with $i = 1$. We write

$$\frac{d^2\bar{r}}{dt^2} + \frac{\mu\bar{r}}{r^3} = \text{grad}\,\hat{R} \quad ; \quad \mu = k^2(m_1+m_0) \tag{15-13}$$

The $\hat{R}$ contains all the perturbations of the other planets. We first assume the motion of the other planets are known, then $\hat{R} = \hat{R}\,(q_1, q_2, q_3, t)$ where $\bar{r} = \bar{i}q_1 + \bar{j}q_2 + \bar{k}q_3$. We then have

$$\frac{dq_i}{dt} = p_i = \frac{\partial H}{\partial p_i} \tag{15-14}$$

$$\frac{dp_i}{dt} = - \frac{\partial H}{\partial q_i} \tag{15-15}$$

$$H = H_o + \hat{R} \tag{15-16}$$

$$H_o = \frac{1}{2} (p_1^2 + p_2^2 + p_3^2) + \frac{\mu}{r}\,, \tag{15-17}$$

Now apply the Hamilton-Jacobi theory by first solving for the case where $\hat{R} = 0$. We are actually just solving the two body central force field problem. This is called the Kepler problem in space.

We want to transform the original variables q_i, p_i into a canonical set α_i, β_i so that if we use

$$S = S\ (q_i,\ \beta_i,\ t);\ p_i = \frac{\partial S}{\partial q_i};\ \alpha_i = - \frac{\partial S}{\partial \beta_i} \tag{15-18}$$

or

$$S = S\ (q_i,\ \alpha_i,\ t);\ p_i = \frac{\partial S}{\partial q_i};\ \beta_i = \frac{\partial S}{\partial \alpha_i} \tag{15-19}$$

where S is a solution of

$$H^1 = H_o\ (q_i,\ \frac{\partial S}{\partial q_i},\ t) + \frac{\partial S}{\partial t} = 0 \tag{15-20}$$

we will obtain the new canonical variables as constants, i.e.,

$$\frac{d\alpha_i}{dt} = - \frac{\partial H^1}{\partial \beta_i} = 0 \qquad \frac{d\beta_i}{dt} = \frac{\partial H^1}{\partial \alpha_i} = 0 \tag{15-21}$$

These α_i, β_i variables are constants for the two body case but the effect of the other planets (now a non-two body problem) is to perturb these "constants," i.e. they make α_i, β_i variables.

For the two body problem we could select any set of coordinates in order to solve the H_0 Hamilton-Jacobi PDE, however, by selecting spherical coordinates the Hamilton-Jacobi PDE will be a separable one.

16. THE KEPLER PROBLEM IN SPACE

We select a spherical polar coordinate system as shown below.

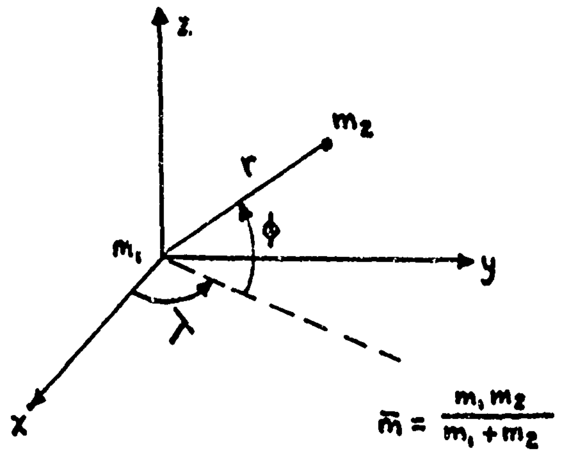


FIGURE 16-1

The polar and rectangular coordinates are related by

$$\begin{aligned} x &= r \cos\phi \cos\lambda \\ y &= r \cos\phi \sin\lambda \\ z &= r \sin\phi \end{aligned} \tag{16-1}$$

Using these coordinates, the kinetic and potential energy becomes

$$T = \frac{1}{2}\bar{m}\,[\dot{r}^2 + r^2 \cos^2\phi\, \dot{\lambda}^2 + r^2\, \dot{\phi}^2] \tag{16-2}$$

$$V = -\frac{\mu}{r} \quad ; \quad \mu = k^2 m_1 m_2 \tag{16-3}$$

Setting $\bar{m} = 1$ and using

$$q_1 = r \qquad q_2 = \phi \qquad q_3 = \lambda \tag{16-4}$$

the Lagrangian becomes

$$L = \frac{1}{2}\,[\dot{q}_1^2 + q_1^2\,\dot{q}_2^2 + q_1^2\,(\cos^2 q_2)\,\dot{q}_3^2] + \frac{\mu}{q_1} \tag{16-5}$$

and the generalized momenta are $\left(p_i = \frac{\partial L}{\partial \dot{q}_i} \right)$

$$p_1 = \dot{q}_1;\; p_2 = q_1^2\,\dot{q}_2;\; p_3 = q_1^2\,(\cos^2 q_2)\,\dot{q}_3 \tag{16-6}$$

There is one ignorable coordinate here, q_3. The Hamiltonian is equal to the total energy.

$$H(p_i, q_i) = \frac{p_1^2}{2} + \frac{p_2^2}{2q_1^2} + \frac{p_3^2}{2q_1^2 \cos^2 q_2} - \frac{\mu}{q_1}\,. \tag{16-7}$$

To solve this problem we must form the Hamilton-Jacobi PDE and solve for the required generating function $S = S\,(q_i, \alpha_i, t)$.

$$H_o + \frac{\partial S}{\partial t} = 0 \qquad p_i = \frac{\partial S}{\partial q_i} \qquad \beta_i = \frac{\partial S}{\partial \alpha_i} \tag{16-8}$$

or

$$\frac{\partial S}{\partial t} + \frac{1}{2}\left(\frac{\partial S}{\partial q_1}\right)^2 + \frac{1}{2q_1^2}\left(\frac{\partial S}{\partial q_2}\right)^2 + \frac{\left(\frac{\partial S}{\partial q_3}\right)^2}{2q_1^2 \cos^2 q_2} - \frac{\mu}{q_1} = 0 \tag{16-9}$$

To solve we assume a separable solution

$$S = S_t(t) + S_1(q_1) + S_2(q_2) + \alpha_3 q_3 \tag{16-10}$$

Since q_3 is ignorable we can represent S as above. Substitution of (16-10) into (16-9) gives

$$\frac{dS_t}{dt} + \frac{1}{2}\left(\frac{dS_1}{dq_1}\right)^2 + \frac{1}{2q_1^2}\left(\frac{dS_2}{dq_2}\right)^2 + \frac{\alpha_3^2}{2q_1^2 \cos^2 q_2} - \frac{\mu}{q_1} = 0 . \tag{16-11}$$

The first term is a function of t alone, the remainder are independent of t. Each of these parts must be a constant, so

$$\frac{dS_t}{dt} = -\alpha_1 \tag{16-12}$$

$$\frac{1}{2}\left(\frac{dS_1}{dq_1}\right)^2 + \frac{1}{2q_1^2}\left(\frac{dS_2}{dq_2}\right)^2 + \frac{\alpha_3^2}{2q_1^2 \cos^2 q_2} - \frac{\mu}{q_1} = \alpha_1 . \tag{16-13}$$

Equation (16-12) may be integrated directly to give

$$S_t = -\alpha_1 t \tag{16-14}$$

dropping the additive constant. If (16-13) is multiplied by $2q_1^2$, it may be separated.

$$\left(\frac{dS_2}{dq_2}\right)^2 + \frac{\alpha_3^2}{\cos^2 q_2} = -q_1^2\left(\frac{dS_1}{dq_1}\right)^2 + 2q_1^2\alpha_1 + 2\mu q_1 \tag{16-15}$$

The left side is a function of S_2 alone, the right of S_1 alone. Therefore both sides must equal some non-negative constant.

$$q_1^2 \left(\frac{dS_1}{dq_1}\right)^2 - 2\alpha_1 q_1^2 - 2\mu q_1 = -\alpha_2^2 \tag{16-16}$$

$$\left(\frac{dS_2}{dq_2}\right)^2 + \frac{\alpha_3^2}{\cos^2 q_2} = \alpha_2^2 \ . \tag{16-17}$$

It may be verified that the solutions to (16-16) and (16-17) are respectively,

$$S_1(q_1) = \pm \int \frac{1}{q_1} \sqrt{2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2} \, dq_1 \tag{16-18}$$

$$S_2(q_2) = \pm \int \sqrt{\alpha_2^2 - \alpha_3^2 \sec^2 q_2} \, dq_2 \tag{16-19}$$

and thus the complete solution of PDE (16-9) is

$$S = -\alpha_1 t + \int \sqrt{2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2} \, \frac{dq_1}{q_1} + \int \sqrt{\alpha_2^2 - \alpha_3^2 \sec^2 q_2} \, dq_2 + \alpha_3 q_3 \tag{16-20}$$

The other constants are found from

$$\beta_i = \frac{\partial S}{\partial \alpha_i} \tag{16-21}$$

or

$$\beta_1 = \frac{\partial S}{\partial \alpha_1} = -t + \int \frac{q_1\, dq_1}{\sqrt{2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2}} \tag{16-22}$$

$$\beta_2 = \frac{\partial S}{\partial \alpha_2} = -\int \frac{\alpha_2\, dq_1}{q_1\sqrt{2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2}} + \alpha_2 \int \frac{dq_2}{\sqrt{\alpha_2^2 - \alpha_3^2 \sec^2 q_2}} \tag{16-23}$$

$$\beta_3 = \frac{\partial S}{\partial \alpha_3} = q_3 - \alpha_3 \int \frac{\sec^2 q_2\, dq_2}{\sqrt{\alpha_2^2 - \alpha_3^2 \sec^2 q_2}} \tag{16-24}$$

If we check the Jacobian of the transformation with respect to the α_i, β_i, it is not identically equal to zero so they are independent parameters (page 115).

We would now like to identify these α_i and β_i in terms of conic parameters. We also need to know what limits to use on the integration. The β's are arbitrary up to this point but we need to fix them for subsequent transformations. We fix the β's to particular β's by selecting the limits of integration and since this amounts to selecting the additive constants for the S function, we have a great deal of freedom in selection. We proceed by applying initial conditions.

From equation (16-22) we can develop the following

$$\frac{d\beta_1}{dt} = -1 + \frac{q_1 \frac{dq_1}{dt}}{\sqrt{2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2}} = 0 \tag{16-25}$$

since β_1 is a constant with respect to time. Here we have made use of

$$\frac{d}{dt}\left[\int_{r_0}^{r} f(r)dr\right] = f(r)\frac{dr}{dt} - f(r_0)\frac{dr_0}{dt} + \int_{r_0}^{r} \frac{\partial f(r)}{\partial t} dr \tag{16-26}$$

where

$$\frac{\partial f(r)}{\partial t} = 0\,; \qquad \frac{dr_o}{dt} = 0.$$

Recall q_1 came from equation (16-16) which we shall shortly discuss further, and q_1 is identified with r, i.e., equation (16-4) says $q_1 = r$ and thus (16-25) is actually

$$r\,\frac{dr}{dt} = \sqrt{2\alpha_1\, r^2 + 2\mu r - \alpha_2^2} \tag{16-27}$$

When $\frac{dr}{dt} = 0$ we are at apogee or perigee for an elliptical orbit. Hence at perigee

$$2\alpha_1\, r_P^2 + 2\mu r_P - \alpha_2^2 = 0\,. \tag{16-28}$$

There are actually two roots to this equation, one is r_P and the other is r_A. From the algebra of quadratic equations we have (r_A and r_P are the two roots of (16-28)).

$$r_A + r_P = -\frac{\mu}{\alpha_1} = 2a \tag{16-29}$$

$$r_A r_P = -\frac{\alpha_2^2}{2\alpha_1} = a^2\,(1-e^2) \tag{16-30}$$

since

$$r_P = a\ (1-e);\ r_A = a\ (1+e), \tag{16-31}$$

From (16-29) we have

$$\alpha_1 = -\frac{\mu}{2a} = E = \text{total energy}. \tag{16-32}$$

Substitution of α_1 and r_P from (16-32) and (16-31) into equation (16-28) gives

$$-\frac{2\mu a^2(1-e)^2}{2a} + 2\mu a(1-e) - \alpha_2^2 = 0 \tag{16-33}$$

or

$$\alpha_2 = \sqrt{\mu a\ (1-e^2)} = h = \text{angular momentum} \tag{16-34}$$

So we decide that r goes from perigee to apogee. We take the initial time $t = \tau_o$ and assume we are initially at perigee. Thus τ_o is the time of passage of perigee. We can now write

$$\beta_1 = -t + \int_{r_p}^{r} \frac{q_1\, dq_1}{\sqrt{2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2}} \tag{16-35}$$

at $r = r_P$, $t = \tau_o$ so the above gives

$$\beta_1 = -\tau_o\ . \tag{16-36}$$

We can then write equation (16-35) as

$$t - \tau_o = \frac{1}{\sqrt{-2\alpha_1}} \int_{r_p}^{r} \frac{q_1 \, dq_1}{\sqrt{(q_1 - r_p)(r_A - q_1)}} \quad . \tag{16-37}$$

Then using

$$\begin{aligned} r_A &= a\ (1+e) \\ r_p &= a\ (1-e) \\ q_1 &= r = a\ (1-e \cos E) \\ dq_1 &= ae \sin E \, dE \end{aligned} \tag{16-38}$$

we obtain

$$\frac{\sqrt{-2\alpha_1}}{a}\left(t - \tau_o\right) = \int_0^E (1 - e \cos E)\, dE \tag{16-39}$$

where

$$E = \cos^{-1}\left[\frac{a-r}{ae}\right]. \tag{16-40}$$

Integration of (16-39) gives

$$E - e \sin E = \frac{1}{a}\sqrt{-2\alpha_1}\ (t - \tau_o) \tag{16-41}$$

Letting $n = \frac{1}{a}\sqrt{-2\alpha_1}$ and calling it the mean angular velocity, (16-41) becomes Kepler's equation. Further since $\alpha_1 = -\frac{\mu}{2a} = E$ this is actually

$$n = \sqrt{\frac{\mu}{a^3}} \tag{16-42}$$

which is indeed the mean angular velocity as advertised. Since $\mu = k^2\ (m_1 + m_2)$, equation (16-42) is actually $n^2a^3 = k^2(m_1 + m_2)$, which verifies the fact that $n = \frac{2\pi}{P}$, the mean angular velocity.

Now let's return to equation (16-16).

$$q_1^2\left(\frac{dS_1}{dq_1}\right)^2 = 2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2 = F(q_1) \geq 0 \tag{16-43}$$

For any given α_1, α_2 this places restrictions on the allowed values of $q_1 = r$. Two cases arise, depending on whether $\alpha_1 > 0$ or $\alpha_1 < 0$. Since α_1 is actually the total energy we know $\alpha_1 > 0$ corresponds to hyperbolic trajectories and $\alpha_1 < 0$ to elliptical orbits. We can show this from equation (16-43). The roots of the equation are given by

$$q_1 = -\frac{\mu}{2\alpha_1} \pm \sqrt{\frac{\mu}{4\alpha_1^2} + \frac{\alpha_2^2}{2\alpha_1}}\ . \tag{16-44}$$

If $\alpha_1 > 0$, there will be one positive real root, and one negative real root. The general shape of $F(q_1)$ is indicated on the next page.

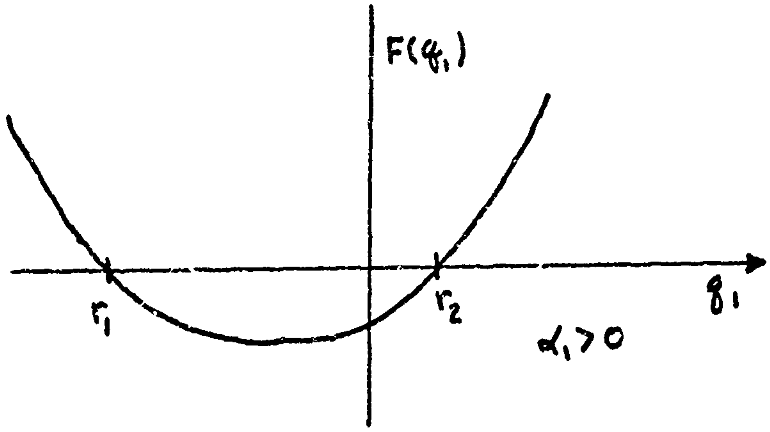


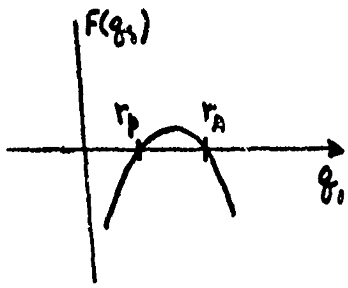
FIGURE 16-2

Since $r < 0$ is not physically meaningful, it is necessary that $q_1 = r > r_2$, which is consistent with a hyperbolic orbit. r_2 is the positive root of (16-44), $\alpha_1 > 0$.

If $\alpha_1 = 0$, equation (16-43) requires $q_1 = r > \frac{\alpha_2^{\,2}}{2\mu} > 0$, which is consistent with a parabolic trajectory.

If $\alpha_1 < 0$, then equation (16-44) indicates two positive roots, call them $r_A > r_p > 0$.

$$r_A = \frac{\mu}{(-2\alpha_1)} + \sqrt{\frac{\mu}{4\alpha_1^2} - \frac{\alpha_2^2}{(-2\alpha_1)}}$$
$$r_p = \frac{\mu}{(-2\alpha_1)} - \sqrt{\frac{\mu}{4\alpha_1^2} - \frac{\alpha_2^2}{(-2\alpha_1)}} \qquad (16\text{-}45)$$



This is consistent with an elliptic trajectory. In all subsequent work, only this case will be considered. If $\alpha_1 < 0$, (16-43) may be written

$$F(q_1) = (-2\alpha_1)\,(q_1 - r_a)\,(-q_1 + r_p) \geq 0. \qquad (16\text{-}46)$$

This will be true for $r_a < q_1 < r_p$. The excursions of r are therefore limited.

In addition, equation (16-17) shows that

$$\left(\frac{dS_2}{dq_2}\right)^2 = \alpha_2^2 - \frac{\alpha_3^2}{\cos^2 q_2} \geq 0 \qquad (16\text{-}47)$$

from which

$$\sin^2 q_2 \leq \frac{\alpha_2^2 - \alpha_3^2}{\alpha_2^2} = 1 - \left(\frac{\alpha_3}{\alpha_2}\right)^2 \leq 1 \qquad (16\text{-}48)$$

This shows that $q_2 = \phi$, the latitude, has limited excusions and that

$$\alpha_3^2 \leq \alpha_2^2 \qquad (16\text{-}49)$$

We will use these latter tid-bits to aid in the evaluation of the other constants.

We now have evaluated α_1, α_2, and β_1 in terms of conic parameters. To proceed further to identify the other canonical constants we select certain parameters, Ω, ω and i from spherical geometry.

Refer to Figure 16-3. A fixed inertial x, y, z system is given. The inclination of the orbit plane (we will shortly prove that it <u>is</u> a plane) with respect to the equatorial x-y plane is i. The line NN^1 where the equatorial x-y plane and the orbital plane intersect is called the line of nodes. The perigee of the orbit is at P and its projection on the celestial sphere is P^1. The particle is at M and its projection at M^1. N is called the ascending node because the particle is moving upward as it crosses N. Correspondingly N^1 is called the descending node. The angle Ω = XON is the longitude of the ascending node or nodal angle.

ω= NOP^1 is the argument of perifocus or for the earth, just the argument of perigee.

The quantities Ω, ω, i serve to locate the orbit plane and the orbit within the plane. The shape of the orbit is determined by a the <u>semi-major axis</u> and e the <u>eccentricity</u>. The position along the orbit may be determined by fixing τ, the <u>time of perigee passage</u>. Thus the six quantities Ω, ω, i a, e τ serve to specify the orbit completely. These are called the <u>orbital elements</u>, and they give a relatively clear physical picture of the orbit.

The principal objective of this section is to develop the relationships between these orbital elements and the constants of integration α_1, α_2, α_3, β_1, β_2, β_3 which arose in the solution of the Kepler problem. Thus far we have identified α_1, α_2, and β_1. Now let's proceed to evaluate the rest.

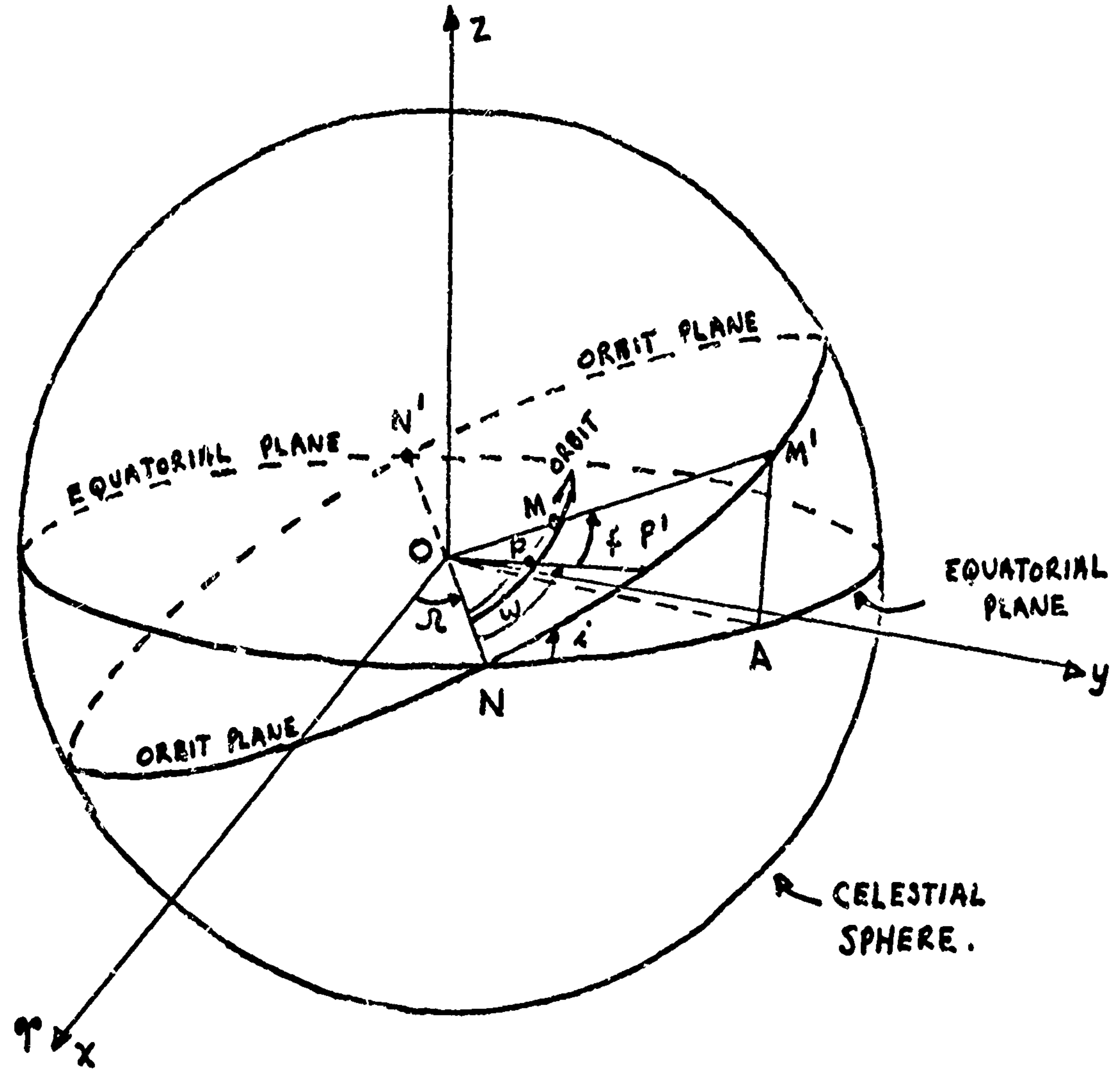


FIGURE 16-3. CELESTIAL SPHERE AND ELEMENTS OF AN ORBIT.

Referring to equation (16-24), it may be written as

$$\beta_3 = \lambda - \int_0^{\phi} \frac{\sec^2 x\,dx}{\sqrt{\frac{\alpha_2^2 - \alpha_3^2}{\alpha_3^2} - \tan^2 x}} \tag{16-50}$$

It has already been established $\alpha_2^2 \geq \alpha_3^2$ (16-49). Accordingly, there exists some positive acute angle A such that

$$\alpha_2^2 - \alpha_3^2 = \alpha_3^2 \tan^2 A \tag{16-51}$$

Using this, the integral of (16-50) becomes,

$$\beta_3 - \lambda = -\int_0^{\phi} \frac{\sec^2 x\,dx}{\sqrt{\tan^2 A - \tan^2 x}} \tag{16-52}$$

Integrating

$$\beta_3 - \lambda = \sin^{-1}\left[-\frac{\tan\phi}{\tan A}\right] \tag{16-53}$$

or

$$\sin(\beta_3 - \lambda) = -\frac{\tan\phi}{\tan A}. \tag{16-54}$$

Since λ is going to go through 360° each orbit, it follows that both sides of this equation must go through the full range from -1 to +1, but not outside this range. This means that

$$\phi_{max} = A, \ \phi_{min} = -A \tag{16-55}$$

Equation (16-54) may be written:

$$\sin\beta_3 \cos\lambda - \cos\beta_3 \sin\lambda = - \tan\phi \ \cot A . \tag{16-56}$$

Multiply both sides by r $\cos\phi$

$$(r \cos\phi \cos\lambda)\sin\beta_3 - (r \cos\phi \sin\lambda)\cos\beta_3 = - (r \sin\phi)\cot A . \tag{16-57}$$

Making use of equations (16-1), this may be expressed as:

$$x \sin\beta_3 - y \cos\beta_3 + z \cot A = 0 . \tag{16-58}$$

This is clearly the equation of a plane through the origin. This would be expected from physical principles, but this is a direct proof that the orbit lies in a plane containing the attracting center. Big deal.

However, this being the case we can then safely say that the maximum value of ϕ is simply the inclination angle i between the x-y plane and the orbit plane, i.e., A = i. Hence using

$$\cos^2 A = \frac{1}{1 + \tan^2 A} \tag{16-59}$$

equation (16-51) gives

$$\alpha_3 = \alpha_2 \cos i = \cos i \sqrt{\mu a(1-e^2)} \qquad (16\text{-}60)$$

When $\phi = 0$, the moving particle lies in the equatorial plane. This means that it lies on the line NN^1. Since

$$\sin(\beta_3 - \lambda) = -\frac{\tan\phi}{\tan i} \qquad (16\text{-}61)$$

we must have

$$\beta_3 - \lambda = k\pi \qquad k = 0, \pm 1, \pm 2, \ldots \qquad (16\text{-}62)$$

β_3 is clearly the longitude of either the ascending or descending node. We choose k so that

$$\beta_3 = \Omega \,. \qquad (16\text{-}63)$$

This is actually evident from (16-52) when $\phi = 0$ then $\beta_3 = \lambda$ where λ is the value at the time $\phi = 0$, i.e., when $\lambda = \Omega$.

We now need only evaluate β_2. Equation (16-23) gives

$$\beta_2 = -\int_{r_p}^{r} \frac{\alpha_2\, dr}{r\sqrt{2\alpha_1 r^2 + 2\mu r - \alpha_2^2}} + \alpha_2 \int_0^{\phi} \frac{d\phi}{\sqrt{\alpha_2^2 - \alpha_3^2 \sec^2\phi}} \qquad (16\text{-}64)$$

$$\beta_2 = -I_1 + I_2 \qquad (16\text{-}65)$$

Let's treat the two integrals separately.

I_1 may be treated by a method similar to that used previously.

$$I_1 = \frac{\alpha_2}{\sqrt{-2\alpha_1}} \int_{r_p}^{r} \frac{dx}{x\sqrt{(x-r_p)(-x+r_A)}} \tag{16-66}$$

Making the change of variables as with equations (16-35) to (16-39) gives (page 155),

$$I_1 = \frac{\alpha_2}{a\sqrt{-2\alpha_1}} \int_0^E \frac{dE}{(1-e\cos E)} \tag{16-67}$$

which, on integration gives:

$$I_1 = 2\tan^{-1}\left[\sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2}\right] \tag{16-68}$$

Which may also be written

$$\tan\frac{I_1}{2} = \sqrt{\frac{1+e}{1-e}}\tan\frac{E}{2}. \tag{16-69}$$

Comparison of this with the well-known* equations for elliptic orbits, it is seen that I_1 is simply the true anomaly f. Refer to equation (3-20) at the beginning of these notes.

Now let's turn to I_2.

$$I_2 = \alpha_2 \int_0^{\phi} \frac{dx}{\sqrt{\alpha_2^2 - \alpha_3^2\sec^2 x}} \tag{16-70}$$

*To those who know it well; for others see equation (3-20).

Using $\sin^2 x = 1 - \cos^2 x$ this may be reduced to

$$I_2 = \int_0^{\phi} \frac{\cos x \, dx}{\sqrt{\frac{\alpha_2^2 - \alpha_3^2}{\alpha_2^2} - \sin^2 x}} \tag{16-71}$$

From equation (16-51) and the fact that A = i, we can write

$$\tan i = \frac{\sqrt{\alpha_2^2 - \alpha_3^2}}{\alpha_3}$$

and hence

$$\sin^2 i = \frac{\alpha_2^2 - \alpha_3^2}{\alpha_2^2} \tag{16-72}$$

so that I_2 can be written as

$$I_2 = \int_0^{\phi} \frac{\cos x \, dx}{\sqrt{\sin^2 i - \sin^2 x}} = \int_0^{\phi} \frac{du}{\sqrt{\sin^2 i - u^2}} \tag{16-73}$$

Integration gives

$$I_2 = \sin^{-1}\left[\frac{\sin\phi}{\sin i}\right] \tag{16-74}$$

Combining I_1 and I_2, equation (16-65) may be written

$$\beta_2 + f = \sin^{-1}\left[\frac{\sin\phi}{\sin i}\right] \tag{16-75}$$

$$\sin(\beta_2 + f) = \frac{\sin\phi}{\sin i} \tag{16-76}$$

Now consider the spherical triangle NAM' of Figure 16-3. The angle $M'AN$ is a right angle. Call $NOM' = \psi$. Then

$$\frac{\sin \phi}{\sin i} = \frac{\sin \psi}{\sin 90°} = \sin \psi \tag{16-77}$$

Comparison with (16-76) shows that $\psi = f + \beta_2$. From definitions given earlier (see Figure 16-3), $\psi = \omega + f$. Therefore

$$\omega = \beta_2 \tag{16-78}$$

which is the final relationship. Collecting them all:

$$\begin{aligned} \alpha_1 &= -\frac{\mu}{2a} = E & \beta_1 &= -\tau_o \\ \alpha_2 &= \sqrt{\mu\, a(1-e^2)} = h & \beta_2 &= \omega \\ \alpha_3 &= \cos i \sqrt{\mu a(1-e^2)} = h \cos i & \beta_3 &= \Omega \end{aligned} \tag{16-79}$$

The elements of the orbit in terms of the canonical constants are

$$\begin{aligned} a &= -\frac{\mu}{2\alpha_1} & \tau_o &= -\beta_1 \\ e &= \sqrt{1 + \frac{2\alpha_1 \alpha_2^2}{\mu^2}} & \omega &= \beta_2 \\ i &= \cos^{-1}\left[\frac{\alpha_3}{\alpha_2}\right] & \Omega &= \beta_3 \end{aligned} \tag{16-80}$$

In this example, we begin to see the power and elegance of the Hamilton-Jacobi method. As a part of the solution, we have proved directly that the orbit lies in a plane and that it is an ellipse with the attracting center at one focus. We have derived Kepler's equations. We have determined the coordinates as functions of time and the initial conditions. The solution was obtained by straightforward application of the method.

Only in interpreting the solution was use made of known properties of elliptic orbits. Even there, this knowledge was used only as a guide in making variable changes in the various integrals.

The required coordinate transformation between q_i, p_i and α_i, β_i are derived from the following:

$$p_i = \frac{\partial S}{\partial q_i} \qquad \beta_i = \frac{\partial S}{\partial \alpha_i} \tag{16-81}$$

$$p_1 = \frac{\partial S}{\partial q_1} = \frac{1}{q_1}\sqrt{2\alpha_1 q_1^2 + 2\mu q_1 - \alpha_2^2}$$

$$p_2 = \frac{\partial S}{\partial q_2} = \sqrt{\alpha_2^2 - \alpha_3^2 \sec^2 q_2} \tag{16-82}$$

$$p_3 = \frac{\partial S}{\partial q_3} = \alpha_3 \tag{16-83}$$

The β_i values are listed in equations (16-22) to (16-23). q_1 may be determined from (16-22). This may be substituted into (16-23) to

determine q_2. This in turn could be substituted into equation (16-23) to determine q_3. These together with equations (16-83) above are sufficient to determine the transformation equations. The student should carry out the necessary integration and substitution.

17. OTHER CANONICAL ELEMENTS

The change of variables from one canonical set α_i, β_i to another is facilitated by using

$$\hat{S} = \Sigma\, \alpha_i\, \beta_i \tag{17-1}$$

where the β_i (or the α_i) are expressed in terms of half of the variables of the new set. For example let's take three of the variables of the new set to be, ℓ, g, h defined by

$$g = \beta_2 \qquad h = \beta_3 \qquad \ell = \frac{\alpha_1}{\mu} \sqrt{-2\alpha_1}\,(t + \beta_1) \tag{17-2}$$

so that

$$\beta_1 = \frac{\mu \ell}{\alpha_1 \sqrt{-2\alpha_1}} - t$$

Then

$$\hat{S} = \frac{\mu}{\sqrt{-2\alpha_1}}\,\ell - \alpha_1 t + \alpha_2 g + \alpha_3 h \tag{17-3}$$

For the other three canonical variables, L, G, H we must have

$$L = \frac{\partial \hat{S}}{\partial \ell} = \mu(-2a_1)^{-\frac{1}{2}} = \sqrt{\mu a} \qquad \ell = n\ (t-\tau_o) = nt + \chi$$

$$G = \frac{\partial \hat{S}}{\partial g} = \alpha_2 = \sqrt{\mu a(1-e^2)} = h = L\ \sqrt{1-e^2} \qquad g = \omega \tag{17-4}$$

$$H = \frac{\partial \hat{S}}{\partial h} = \alpha_3 = h\ \cos_i = G\ \cos_i \qquad h = \Omega$$

and furthermore

$$H^1 = H + \frac{\partial \hat{S}}{\partial t} = H - \alpha_1 = H + \frac{\mu^2}{2L^2} \tag{17-5}$$

These ℓ, g, h and L, G, H are the Delaunay canonical elements that are of such importance in celestial mechanics that one uses the symbol F for the Hamiltonian and reserves H for the above Delauney canonical variables. Even worse, celestial mechanics uses Hamiltonian = -F.

One can also use a set of modified Delaunay canonical variables.

$$L^1 = L = \sqrt{\mu a} \qquad \ell^1 = \ell + g + h = nt + \omega - n\tau_o$$

$$G^1 = G - L = \sqrt{\mu a}\ (\sqrt{1-e^2}-1) \qquad g^1 = g + h = \omega + \Omega \tag{17-6}$$

$$H^1 = H - G = (\cos i - 1)\ \sqrt{\mu a(1-e^2)} \qquad h^1 = h = \Omega$$

Another set is Poincare's canonical variables.

$$L = \sqrt{\mu a} \qquad \lambda = \omega + \Omega + \chi$$

$$\xi_1 = \rho\ \sin(\omega+\Omega) \qquad \eta_1 = \rho\ \cos(\omega+\Omega)$$

$$\xi_2 = \sigma\ \sin\ \Omega \qquad \eta_2 = \sigma\ \cos\ \Omega$$

where

$$\rho = \sqrt{2L(1-\sqrt{1-e^2})}$$

$$\sigma = \sqrt{2L(1-\cos i)\sqrt{1-e^2}} \tag{17-8}$$

The Poincaré elements can be obtained from the modified Delaunay elements by means of the transformation

$$q_j^1 = \sqrt{2q_j} \cos p_j$$

$$p_j^1 = \sqrt{2q_j} \sin p_j \ . \tag{17-9}$$

18. ELLIPTICAL VARIABLES

The orbital elements, because of their geometrical significance, are more often used than the canonical variables. This complicates the form of the equations slightly. With the canonical variables we have

$$\frac{d\alpha_i}{dt} = -\frac{\partial \hat{R}}{\partial \beta_i}; \qquad \frac{d\beta_i}{dt} = \frac{\partial \hat{R}}{\partial \alpha_i} \tag{18-1}$$

But celestial mechanics reverses the sign of the disturbing function. One uses $R = -\hat{R}$. This in turn is related to a different definition for the potential function. In physics one usually writes

$$\frac{d^2\bar{r}}{dt^2} = \nabla V \tag{18-2}$$

where V is the potential, for example, for the two body case $V = -\frac{\mu}{r}$. In celestial mechanics one uses instead

$$\frac{d^2\bar{r}}{dt^2} = -\nabla U \tag{18-3}$$

where U is the potential which for the two body case is $U = \frac{\mu}{r}$. Thus

$$U = -V \tag{18-4}$$

Now refer to equations (15-1) to (15-10). Instead of V we now write in terms of U.

$$H = T + V = T - U = T - U_o = R = H_o - R \tag{18-5}$$

and hence $\hat{R} = -R$ and (18-1) becomes

$$\frac{d\alpha_i}{dt} = + \frac{\partial R}{\partial \beta_i}; \qquad \frac{d\beta_i}{dt} = - \frac{\partial R}{\partial \alpha_i} \tag{18-6}$$

Thus the hatted function $\hat{R}$ corresponds to a potential V, defined by equation (18-2) and R corresponds to U, a potential defined by equation (18-3). <u>We shall hereinafter always use the unhatted version and will be speaking of the potential as U</u>.

We thus have

$$\frac{d\alpha_i}{dt} = \frac{\partial R}{\partial \beta_i} \qquad \frac{d\beta_i}{dt} = - \frac{\partial R}{\partial \alpha_i} \tag{18-6}$$

To solve we thus require $R = R(\alpha_i, \beta_i, t)$. To change from canonical variables, α_i and β_i, to the orbital elements, we denote these elements by a_m (m = 1, 2, 3, 4, 5, 6). (For example $a_1 = a$ $a_2 = e$ etc.) then we can write

$$\frac{da_m}{dt} = \sum_{r=1}^{3} \left(\frac{\partial a_m}{\partial \alpha_r} \frac{d\alpha_r}{dt} + \frac{\partial a_m}{\partial \beta_r} \frac{d\beta_r}{dt} \right) \tag{18-7}$$

Using (18-6) this is

$$\frac{da_m}{dt} = \sum_{r=1}^{3} \left(\frac{\partial a_m}{\partial \alpha_r} \frac{\partial R}{\partial \beta_r} - \frac{\partial a_m}{\partial \beta_r} \frac{\partial R}{\partial \alpha_r} \right) \tag{18-8}$$

But -

$$\frac{\partial R}{\partial \alpha_r} = \sum_{s=1}^{6} \frac{\partial R}{\partial a_s}\frac{\partial a_s}{\partial \alpha_r} \quad ; \quad \frac{\partial R}{\partial \beta_r} = \sum_{s=1}^{6} \frac{\partial R}{\partial a_s}\frac{\partial a_s}{\partial \beta_r} \quad . \tag{18-9}$$

The operations of (18-8) and (18-9) are lengthly but not difficult. Using (18-9) in (18-8) gives

$$\frac{da_m}{dt} = -\sum_{s=1}^{6} \frac{\partial R}{\partial a_s} \sum_{r=1}^{3} \left(\frac{\partial a_m}{\partial \beta_r}\frac{\partial a_s}{\partial \alpha_r} - \frac{\partial a_m}{\partial \alpha_r}\frac{\partial a_s}{\partial \beta_r} \right) \tag{18-10}$$

This last sum is called the Poisson bracket and is written

$$\left[a_m, a_s \right] = \sum_{r=1}^{3} \left(\frac{\partial a_m}{\partial \beta_r}\frac{\partial a_s}{\partial \alpha_r} - \frac{\partial a_m}{\partial \alpha_r}\frac{\partial a_s}{\partial \beta_r} \right) \tag{18-11}$$

These Poisson brackets are very important in Physics. The student should carefully review Goldstein, pages 250-258, where these bracket relationships are discussed. In particular one can easily show that

$$\left[a_m, a_s \right] = -\left[a_s, a_m \right] \tag{18-12}$$

Using this bracket notation, equation (18-10) can be written as

$$\frac{da_m}{dt} = -\sum_{s=1}^{6} \left[a_m, a_s \right] \frac{\partial R}{\partial a_s} \quad . \tag{18-13}$$

The Poisson brackets are zero for most combinations of a_m and a_s. Using equations (16-80) the results shown in the array below can be obtained.

$$
\begin{array}{c|cccccc}
 & a & e & i & \Omega & \omega & \tau_o \\
\hline
a & 0 & 0 & 0 & 0 & 0 & [a,\tau_o] \\
e & 0 & 0 & 0 & 0 & [e,\omega] & [e,\tau_o] \\
i & 0 & 0 & 0 & [i,\Omega] & [i,\omega] & 0 \\
\Omega & 0 & 0 & [\Omega,i] & 0 & 0 & 0 \\
\omega & 0 & [\omega,e] & [\omega,i] & 0 & 0 & 0 \\
\tau_o & [\tau_o,a] & [\tau_o,e] & 0 & 0 & 0 & 0
\end{array} \qquad (18\text{-}14)
$$

The non-zero brackets, expressed in terms of orbital elements are

$$[\tau_o,a] = -[a,\tau_o] = -\frac{2a^2}{\mu}$$

$$[\omega,e] = -[e,\omega] = -\frac{1}{\mu ae}\sqrt{\mu a(1-e^2)}$$

$$[\tau_o,e] = -[e,\tau_o] = -\frac{a}{\mu e}(1-e^2)$$

$$[\Omega,i] = -[i,\Omega] = -\frac{1}{\sin i\sqrt{\mu a(1-e^2)}}$$

$$[\omega,i] = -[i,\omega] = +\frac{\cot i}{\sqrt{\mu a(1-e^2)}} \quad .$$

Substituting these into equations (18-13), the equations of motion in terms of the orbit elements become,

$$\frac{da}{dt} = -\frac{2a^2}{\mu}\frac{\partial R}{\partial \tau_0}$$

$$\frac{de}{dt} = -\frac{\sqrt{1-e^2}}{e\sqrt{\mu a}}\frac{\partial R}{\partial \omega} - \frac{a(1-e^2)}{\mu e}\frac{\partial R}{\partial \tau_0}$$

$$\frac{di}{dt} = -\frac{1}{\sqrt{\mu a(1-e^2)}\,\sin i}\frac{\partial R}{\partial \Omega} + \frac{\cot i}{\sqrt{\mu a(1-e^2)}}\frac{\partial R}{\partial \omega}$$

$$\frac{d\Omega}{dt} = \frac{1}{\sqrt{\mu a(1-e^2)}\,\sin i}\frac{\partial R}{\partial i} \tag{18-16}$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{e\sqrt{\mu a}}\frac{\partial R}{\partial e} - \frac{\cot i}{\sqrt{\mu a(1-e^2)}}\frac{\partial R}{\partial i}$$

$$\frac{d\tau_0}{dt} = \frac{2a^2}{\mu}\frac{\partial R}{\partial a} + \frac{a(1-e^2)}{\mu e}\frac{\partial R}{\partial e}$$

In order to obtain a solution one must be able to express R as a function of the orbital elements. This may be difficult or even nearly impossible.

$$R = R\,(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, t)$$

$$R = R\,(a_1, a_2, a_3, a_4, a_5, a_6, t) \tag{18-17}$$

$$R = R\,(a, e, \omega, \Omega, i, \tau_o, t)$$

Sometimes one uses $\chi = -n\tau_o$ as a variable in place of τ_o. The new equations may be found as before by finding the corresponding new Poisson brackets. This result gives.

$$\frac{da}{dt} = 2\sqrt{\frac{a}{\mu}}\ \frac{\partial R}{\partial \chi}$$

$$\frac{de}{dt} = \frac{1-e^2}{e\sqrt{\mu a}}\frac{\partial R}{\partial \chi} - \frac{\sqrt{1-e^2}}{e\sqrt{\mu a}}\frac{\partial R}{\partial \omega}$$

$$\frac{d\Omega}{dt} = \frac{1}{\sqrt{\mu a(1-e^2)}\ \sin i}\frac{\partial R}{\partial i}$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{e\sqrt{\mu a}}\frac{\partial R}{\partial e} - \frac{\cot i}{\sqrt{\mu a(1-e^2)}}\frac{\partial R}{\partial i} \qquad (18\text{-}18)$$

$$\frac{di}{dt} = \frac{\cot i}{\sqrt{\mu a(1-e^2)}}\left[\frac{\partial R}{\partial \omega} - \frac{1}{\cos i}\frac{\partial R}{\partial \Omega}\right]$$

$$\frac{d\chi}{dt} = -\frac{1}{\sqrt{\mu a}}\left[\frac{1-e^2}{e}\frac{\partial R}{\partial e} + 2a\frac{\partial R}{\partial a}\right]$$

$$n = \frac{\sqrt{k^2(M+m)}}{a^{3/2}} \qquad \chi = -n\tau_o \qquad \chi = -\sqrt{\frac{\mu}{a^3}}\,\tau_o$$

$$\mu = k^2(M+m)$$

We shall refer to these equations many times in subsequent developments. We can also write the equations (18-18) by replacing $\mu = n^2a^3$ or $\sqrt{\mu a} = na^2$.

$$\frac{da}{dt} = \frac{2}{na}\frac{\partial R}{\partial \chi}$$

$$\frac{de}{dt} = \frac{1-e^2}{na^2e}\frac{\partial R}{\partial \chi} - \frac{\sqrt{1-e^2}}{na^2e}\frac{\partial R}{\partial \omega}$$

$$\frac{d\Omega}{dt} = \frac{1}{na^2\sqrt{1-e^2}\,\sin i}\frac{\partial R}{\partial i}$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{na^2e}\frac{\partial R}{\partial e} - \frac{\cot i}{na^2\sqrt{1-e^2}}\frac{\partial R}{\partial i} \tag{18-19}$$

$$\frac{di}{dt} = \frac{\cot i}{na^2\sqrt{1-e^2}}\left[\frac{\partial R}{\partial \omega} - \frac{1}{\cos i}\frac{\partial R}{\partial \Omega}\right]$$

$$\frac{d\chi}{dt} = -\frac{1}{na^2}\left[\frac{1-e^2}{e}\frac{\partial R}{\partial e} + 2a\frac{\partial R}{\partial a}\right]$$

$$n = \frac{\sqrt{k^2(M+m)}}{a^{3/2}}$$

Note that equations 18-16, 18-18 and 18-19 are <u>exact</u> - no approximations have been made in deriving them. We started with Newton's laws

changed variables. This particular form of these equations is called the Lagrange planetary equations.

There is an important concept associated with these planetary equations. These elements are called osculating elements. In the case when $R = 0$, i.e., no perturbations, the particle moving in a pure inverse square central force field will move on an ellipse. Unfortuantely, through a given point in space we can erect an infinity of ellipses. However, at any given instant, the particle will have not only a position but a unique velocity as well. There exists a unique ellipse, based on this unperturbed force field which, at every instant in question, has the same position and velocity as the particle moving under the influence of the perturbation. These elements are called the osculating ellipse. Since the osculating ellipse is based on the instantaneous values of position and velocity, it is sometimes called the "instantaneous" ellipse. However, the antiquarian charm of "osculating" causes the term to persist throughout all of mathematics.

Another way of understanding this idea is to consider a particle moving in a perturbed inverse square force field. If at some instant the perturbations are removed ($R \equiv 0$) and only the inverse square field remains, the particle will move on an ellipse and that ellipse is the osculating ellipse.

The elliptical elements of equation (18-18) are not constant when $R \neq 0$, they are variables. The solution to (18-18) in the form

$$\mathfrak{a}_m = \mathfrak{a}_m(t)$$

are the osculating elements at each instant of time.

19. OVERSIMPLIFIED MODEL TO SHOW TECHNIQUE

In order to show how the Lagrange planetary equations (18-18) are solved, let us take a very simple model. Assume two planets are moving in circles, in a plane, about the sun.

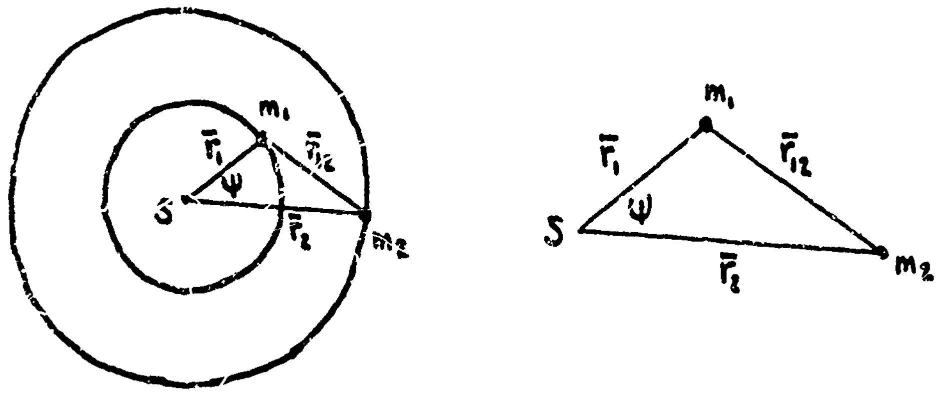


FIGURE 19-1

$$R = -k^2 m_2 \left(\frac{1}{r_{12}} - \frac{\bar{r}_1 \cdot \bar{r}_2}{r_2^3} \right) \tag{19-1}$$

For planet m_1: $r_1 = a$, $r_2 = a_2$ $e_1 = e_2 = 0$

$M = n(t-\tau_o) = E = f$ for each planet.

From Figure 19-1,

$$\psi = f_1 - f_2 = n_1 t - n_2 t - [f_1(0) - f_2(0)] \tag{19-2}$$

where f is a linear function of time for circular orbits. f(0) refers to the value of the true anomaly at time $t = \tau_0$, some arbitrarily selected initial time.

We then develop R as a function of the orbital parameters. From equation (19-1)

$$R = -k^2 m_2 \left[\frac{1}{\sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\psi}} - \frac{r_1 r_2 \cos\psi}{r_2^3} \right] \tag{19-3}$$

$$R = -k^2 m_2 \left[\frac{1}{\sqrt{a_1^2 + a_2^2 - 2a_1 a_2 \cos\psi}} - \frac{a_1 \cos\psi}{a_2^2} \right] \tag{19-4}$$

$$R = -\frac{k^2 m_2}{a_2} \left[\frac{1}{\sqrt{1 + \left(\frac{a_1}{a_2}\right)^2 - 2\frac{a_1}{a_2}\cos\psi}} - \frac{a_1}{a_2}\cos\psi \right] \tag{19-5}$$

Let $\delta = \frac{a_1}{a_2}$. $0 < \delta < 1$ so we can expand the first term of (19-5) in a series.

$$R = -\frac{k^2 m_2}{a_2} \left[1 - \frac{1}{2}\left(\delta^2 - 2\delta\cos\psi\right) + \frac{3}{8}\left(\delta^2 - 2\delta\cos\psi\right)^2 + \ldots\ldots - \delta\cos\psi \right] \tag{19-6}$$

From this process we will eventually be able to reduce to a Fourier series by properly collecting terms.

$$R = -\sum_{k=0}^{\infty} \hat{C}_k \cos^k \psi \tag{19-7}$$

where

$$\psi = n_1 t - n_2 t - n_1(\tau_o)_1 - n_2(\tau_c)_2$$

and $\hat{C}_k$ are power series in δ. This infinite power series can, in turn, be converted to an infinite Fourier series of the form

$$R = \sum_{k=0}^{\infty} C_k \cos k\,\psi \tag{19-8}$$

where C_k is a different power series in δ.

Now consider the variations in the semi-major axis, from equations (18-16) we have

$$\frac{da_1}{dt} = -\frac{2a_1^2}{\mu}\,\frac{\partial R}{\partial(\tau_o)_1}\ . \tag{19-9}$$

From our series expansion for R

$$\frac{\partial R}{\partial(\tau_o)_1} = -\sum_{k=0}^{\infty} k\,(C_k \sin k\,\psi)\,n_1 \tag{19-10}$$

therefore

$$\frac{da_1}{dt} = \frac{2a_1^2}{\mu}\sum_{k=0}^{\infty} n_1 k\,C_k \sin k\,\psi \tag{19-11}$$

where $\psi = (n_1 - n_2)\,t + \text{constant}$. We can integrate (19-11) term by term to give

$$a_1 = \text{constant} + \sum_{k=0}^{\infty} \frac{2n_1 \; C_k \; a_1^2 \cos k\psi}{\mu(n_1-n_2)} \tag{19-12}$$

This is the initial perturbed value of a_1. We now have $a_1 = f(t)$ and we can find similar developments for the other elliptical elements. We could then use this new a_1 in a similar development for a_2 to get a better approximation to a_2; then using this improved value of $a_2 = g(t)$, the process above for finding a_1 could be repeated with $a_2 = g(t)$. The process then goes on and on. The first essential step was in the development of a meaningful series expansion for R. To do this we needed a small parameter.

Note that in integrating equation (19-11) we tacitly assumed the elliptic parameters on the right hand side were practically constant over the integration interval. To be more accurate we should consider not just

$$\frac{da_m}{dt} = F_m\,(a_i, t) \tag{19-13}$$

but should expand this into a power series of deviations from the mean orbital elements (the ones we considered practically constant in the integration of (19-11) to obtain (19-12)). We thus have

$$\frac{da_i}{dt} = \left(F_i\right)_0 + \sum_j \left(\frac{\partial F_i}{\partial a_j}\right)_0 d_1 a_j + \frac{1}{2} \sum_i \sum_j \left(\frac{\partial^2 F_i}{\partial a_i \partial a_j}\right)_0 d_1 a_i \, d_1 a_j + \cdots \tag{19-14}$$

Integration can then be carried out by parts, i.e.,

$$a_m - (a_m)_o = \int (F_i)_o \, dt + \sum_i \left[G_{ij} \, d_1 a_j \right] - \sum_j \int G_{ij} \frac{da_j}{dt} \, dt \tag{19-15}$$

where

$$G_{ij} = \int \left(\frac{\partial F_i}{\partial a_j} \right)_o dt \, . \tag{19-16}$$

Use of this equation entails considerable blood, sweat and tears but it is straightforward. $(F_i)_o$ means the F_i function, which corresponds to the right sides of equations (18-18), which contains only the osculating elements. $\left(\frac{\partial F_i}{\partial a_j}\right)_o$ indicates that after forming the partial derivatives, it is evaluated for the osculating elements. It is multiplied by the appropriate series already obtained in the solution to the first order $d_1 a_j$; the summation indicates all such products are included. Finally the higher order cross products of the first order series enter as well, etc., ...

If we consider the zero, first, second order, etc., developments we can write

$$\frac{da_o}{dt} = 0 \tag{19-17}$$

$$\frac{d[d_1 a]}{dt} = (F_i)_o \tag{19-18}$$

$$\frac{d[d_2 a]}{dt} = \sum_j \left(\frac{\partial F_1}{\partial a_j} \right) d_1 a_j \qquad (19\text{-}19)$$

etc. (lots of luck.)

To carry the problem further in detail one needs more tools for series expansions. This development of the disturbing function in terms of elliptical elements can become a real mess. Most celestial mechanics text books devote two or more chapters to the subject. We shall postpone discussion of these problems until later. First let us consider some modifications of the Lagrange planetary equations (18-18).

20. MODIFICATIONS TO LAGRANGE PLANETARY EQUATIONS

The Lagrange planetary equations (18-18) or (18-19) have a serious defect when computations are made. The disturbing function R is a function of position of the particle and possibly time, i.e., $R = R(x,y,z,t)$. If x, y and z are to be expressed in terms of the orbital elements, it will be necessary to make use of t in this transformation.

$$\begin{aligned} x &= r\,[\cos(\omega+f)\cos\Omega - \sin(\omega+f)\cos i\sin\Omega] \\ y &= r\,[\cos(\omega+f)\sin\Omega + \sin(\omega+f)\cos i\cos\Omega] \\ z &= r\,[\sin i\sin(\omega+f)] \end{aligned} \tag{20-1}$$

$$r = a\,(1 - e\cos E) = \frac{a(1-e^2)}{1 + e\cos f} \tag{20-2}$$

$$E - e\sin E = M = nt + \chi \tag{20-3}$$

$$\tan\frac{f}{2} = \sqrt{\frac{1+e}{1-e}}\,\tan\frac{E}{2} \tag{20-4}$$

In particular, t appears through the mean anomaly M, as

$$M = nt - n\tau_o = nt + \chi \tag{20-5}$$

Since $x = x(a_s,M)$; $y = y(a_s,M)$; $z = z(a_s,M)$, then R may be expressed as $R(a_s,M,t)$, where a_s refers to the six orbital elements.

Since $n^2a^3 = \mu$ we have $n = a^{-3/2}\sqrt{\mu}$ so R depends on the semi-major axis, a, in <u>two ways</u>; <u>explicitly</u>, since a is one of the a_s orbital elements,

and implicitly, through the dependence of M on a. Hence

$$\frac{\partial R}{\partial a} = \left(\frac{\partial R}{\partial a}\right)_M + \left(\frac{\partial R}{\partial M}\right)_a \frac{\partial M}{\partial a} \tag{20-6}$$

where $\left(\frac{\partial R}{\partial a}\right)_M$ denotes the partial derivative with M held constant, (the explicit dependence) and $\left(\frac{\partial R}{\partial M}\right)_a$ the derivative with 'a' held constant, (the implicit dependence). We could live with this arrangement but it can cause trouble and it is easily removed.

First note that the only way R can depend on χ is through the mean anomaly $M = nt + \chi$ and hence

$$\left(\frac{\partial R}{\partial M}\right)_a = \frac{\partial R}{\partial \chi} . \tag{20-7}$$

In addition we have $\frac{\partial M}{\partial a} = t \frac{\partial n}{\partial a}$ but $n = \sqrt{\mu}\, a^{-3/2}$ so that

$$\frac{\partial n}{\partial a} = - \frac{3n}{2a} \text{ or } \frac{\partial M}{\partial a} = - \frac{3nt}{2a} \tag{20-8}$$

and we can write

$$\frac{\partial R}{\partial a} = \left(\frac{\partial R}{\partial a}\right)_M - \frac{3nt}{2a}\left(\frac{\partial R}{\partial \chi}\right) , \tag{20-9}$$

Substituting this into the Lagrange planetary equations for $\dot{\chi}$ we obtain

$$\frac{d\chi}{dt} = -\frac{1-e^2}{na^2e}\frac{\partial R}{\partial e} - \frac{2}{na}\frac{\partial R}{\partial a} \tag{20-10}$$

$$\frac{d\chi}{dt} = -\frac{1-e^2}{na^2e}\frac{\partial R}{\partial e} - \frac{2}{na}\left[\left(\frac{\partial R}{\partial a}\right)_M - \frac{3n}{2a}\frac{\partial R}{\partial \chi}t\right] \tag{20-11}$$

The presence of t here leads to secular terms when the $\dot{\chi}$ equation is integrated. We want to avoid this. From the $\frac{da}{dt}$ equation we have

$$\frac{da}{dt} = \frac{2}{na}\frac{\partial R}{\partial \chi}$$

(see page 178, Equation 17-19)

hence,

$$\frac{d\chi}{dt} = -\frac{(1-e^2)}{na^2e}\frac{\partial R}{\partial e} - \frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_M + \frac{3t}{a^2}\frac{na}{2}\frac{da}{dt} \tag{20-11}$$

$$\frac{d\chi}{dt} - \frac{3n}{2a}\frac{da}{dt}t = \dot{\chi} + \dot{n}t = -\frac{1-e^2}{na^2e}\frac{\partial R}{\partial e} - \frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_M \tag{20-12}$$

The left side is set equal to a new derivative which we call $\dot{\chi}_1$

$$\dot{\chi}_1 = \dot{\chi} + \dot{n}t = \dot{\chi} + \dot{n}t + n - n = \dot{\chi} + \frac{d}{dt}(nt) - n \tag{20-13}$$

Integration of this equation gives

$$\chi_1 = \chi + nt - \int_0^t n\,dt = M - \int_0^t n\,dt \tag{20-14}$$

Now since

$$\frac{\partial \chi_1}{\partial \chi} = 1; \ \frac{\partial R}{\partial \chi} = \frac{\partial R}{\partial \chi_1} \tag{20-15}$$

equation (20-12) may be written

$$\frac{d\chi_1}{dt} = -\frac{1-e^2}{na^2e}\frac{\partial R}{\partial e} - \frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_M . \tag{20-16}$$

This equation (20-16) has the same form as equation (20-10) except that $\left(\frac{\partial R}{\partial a}\right)_M$ appears instead of $\frac{\partial R}{\partial a}$, i.e., we need consider only the explicit derivative. Hence for the Lagrange planetary equations we merely replace χ by χ_1 and use the restricted $\left(\frac{\partial R}{\partial a}\right)_M$. The implicit dependence on 'a' has been absorbed into χ_1

$$M = \chi_1 + \int_0^t n\,dt = E - e \sin E . \tag{20-17}$$

The feasibility of this device is one of the principal reasons for dealing with χ instead of τ_o.

This device also works with other variables which are often used in astronomy.

$$\tilde{\omega} = \omega + \Omega . \tag{20-18}$$

$\tilde{\omega}$ is called the longitude of the pericenter. The true longitude of an object about its primary is given by

$$L = \tilde{\omega} + f = \Omega + \omega + f \tag{20-19}$$

The mean longitude is given by

$$\ell = \tilde{\omega} + n(t - \tau_o) \tag{20-20}$$

$$\ell = \tilde{\omega} + nt + \chi \tag{20-21}$$

and in addition we have the mean longitude at the epoch (ϵ) which is used in place of τ_o or χ.

$$\epsilon = \tilde{\omega} - n\tau_o \tag{20-22}$$

This ϵ is the mean longitude (ℓ) at the moment from which time (t) is reckoned, i.e., the value of ℓ at $t = 0$.

$$\ell = \tilde{\omega} + nt - n\tau_o \tag{20-23}$$

$$\epsilon = \tilde{\omega} - n\tau_o = \omega + \Omega + \chi \tag{20-24}$$

and hence

$$\ell = nt + \epsilon . \tag{20-25}$$

Note that

$$M = \ell - \tilde{\omega} = nt + \epsilon - \tilde{\omega} . \tag{20-26}$$

We can derive the Lagrange planetary equations using the variables $(a, e, i, \Omega, \tilde{\omega}, \epsilon)$ instead of $(a, e, i, \Omega, \omega, \chi)$. Computing the new Poisson brackets and evaluating as before we obtain the following:

$$\frac{da}{dt} = \frac{2}{na}\frac{\partial R}{\partial \epsilon} \tag{20-27}$$

$$\frac{d\epsilon}{dt} = -\frac{2}{na}\frac{\partial R}{\partial a} + \frac{\sqrt{1-e^2}}{\beta n a^2}\frac{\partial R}{\partial e} + \frac{\tan(\frac{i}{2})}{na^2\sqrt{1-e^2}}\frac{\partial R}{\partial i} \tag{20-28}$$

$$\frac{de}{dt} = -\frac{\sqrt{1-e^2}}{ena^2}\frac{\partial R}{\partial \hat{\omega}} - \frac{\sqrt{1-e^2}}{\beta n a^2}\frac{\partial R}{\partial \epsilon} \tag{20-29}$$

$$\frac{d\Omega}{dt} = \frac{1}{na^2 \sin i \sqrt{1-e^2}}\frac{\partial R}{\partial i} \tag{20-30}$$

$$\frac{d\hat{\omega}}{dt} = \frac{\sqrt{1-e^2}}{ena^2}\frac{\partial R}{\partial e} + \frac{\tan(\frac{i}{2})}{na^2\sqrt{1-e^2}}\frac{\partial R}{\partial i} \tag{20-31}$$

$$\frac{di}{dt} = -\frac{1}{na^2 \sin i \sqrt{1-e^2}}\frac{\partial R}{\partial \Omega} - \frac{\tan(\frac{i}{2})}{na^2\sqrt{1-e^2}}\left(\frac{\partial R}{\partial \hat{\omega}} + \frac{\partial R}{\partial \epsilon}\right) \tag{20-32}$$

where

$$\beta = \frac{1+\sqrt{1-e^2}}{e} \quad \text{and} \quad \tan\frac{i}{2} = \frac{\sin i}{1+\cos i} = \frac{1-\cos i}{\sin i} \tag{20-33}$$

Again we have the implicit/explicit dependence of R on 'a' because we have replaced M by ϵ, i.e., $R = R\,(a_s, M, t)$. As before the problem

only occurs in the $\frac{d\epsilon}{dt}$ equation.

Following as before we have

$$-\frac{2}{na}\frac{\partial R}{\partial a} = -\frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_M - \frac{2}{na}\left(\frac{\partial R}{\partial M}\right)_a \frac{\partial M}{\partial a} \qquad (20\text{-}34)$$

$$-\frac{2}{na}\frac{\partial R}{\partial a} = -\frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_n - \frac{2}{na}\left(\frac{\partial R}{\partial \epsilon}\right)_a t\frac{dn}{da} \qquad (20\text{-}34)$$

where again the subscript on the bracket denotes a variable that is constant for that operation. Upon substituting

$$\frac{\partial R}{\partial \epsilon} = \frac{na}{2}\frac{da}{dt}$$ (see page 191, Equation 20-27)

we have

$$-\frac{2}{na}\frac{\partial R}{\partial a} = -\frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_n - t\frac{dn}{da}\frac{da}{dt} \qquad (20\text{-}35)$$

$$-\frac{2}{na}\frac{\partial R}{\partial a} = -\frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_n - t\frac{dn}{dt} \qquad (20\text{-}36)$$

Now substitute this expression for the appropriate term in equation (20-28) give

$$\frac{d\epsilon'}{dt} = \frac{d\epsilon}{dt} + t\frac{dn}{dt} = -\frac{2}{na}\left(\frac{\partial R}{\partial a}\right)_n + \frac{\sqrt{1-e^2}}{\beta na^2}\frac{\partial R}{\partial e} + \frac{\tan\frac{i}{2}}{na^2\sqrt{1-e^2}}\frac{\partial R}{\partial i} \qquad (20\text{-}37)$$

where we have defined a new variable ϵ' as

$$\epsilon' = \epsilon + nt - \int n\,dt \tag{20-38}$$

which integrates to give

$$\epsilon' = \epsilon + nt - \int ndt \ , \tag{20-39}$$

Thus the device works here just as it did in the χ case.

The factor n that occurs in the mean anomaly does not vary when the derivatives are formed with respect to 'a' and we regard the explicit dependence only. Further, nt is replaced by $\int n\,dt$ in the expression for the mean anomaly. The value of n is now found from the 'a' variation by the relation

$$\rho = \int ndt = \sqrt{\mu}\int a^{-3/2}\,dt \tag{20-40}$$

Differentiating this twice with respect to time gives

$$\frac{d^2\rho}{dt^2} = -\frac{3\sqrt{\mu}}{2\,a^{5/2}}\,\frac{da}{dt} = -\frac{3\sqrt{\mu}}{2\sqrt{a^5}}\left(\frac{2}{na}\,\frac{\partial R}{\partial \epsilon}\right) \tag{20-41}$$

$$\frac{d^2\rho}{dt^2} = -\frac{3}{a^2}\,\frac{\partial R}{\partial \epsilon}\ . \qquad \left(\epsilon' = \epsilon + nt - \rho\right) \tag{20-42}$$

The integral $\rho = \int ndt$ is called the mean motion in the disturbed orbit.

— # —

In addition to this "secular" problem, the normal forms of the Lagrange planetary equations, including those of equation (20-27) to

(20-32), are not suited for undisturbed orbits of either small eccentricity or small inclination angles, i.e., for near circular or near equatorial orbits.

The trouble lies in the presence of both e and sin i in the denominators of the right hand sides. In addition, for circular orbits, ω loses its significance and likewise Ω loses its meaning for equatorial orbits. We thus look for slightly different orbital elements for these cases.

For small eccentricity (near circular orbits) we replace e and $\tilde{\omega}$ by

$$h = e \sin \tilde{\omega} \text{ and } k = e \cos \tilde{\omega}$$

where

$\tilde{\omega} = \omega + \Omega$. Then equations (20-29) and (20-31) for $\frac{de}{dt}$ and $\frac{d\tilde{\omega}}{dt}$ are replaced by

$$\frac{dh}{dt} = \frac{\sqrt{1-e^2}}{na^2}\frac{\partial R}{\partial k} - \frac{h\sqrt{1-e^2}}{na^2\left[1+\sqrt{1-e^2}\right]}\frac{\partial R}{\partial \epsilon} + \frac{k\tan(\frac{i}{2})}{na^2\sqrt{1-e^2}}\frac{\partial R}{\partial i} \qquad (20\text{-}43)$$

$$\frac{dk}{dt} = -\frac{\sqrt{1-e^2}}{na^2}\frac{\partial R}{\partial h} - \frac{k\sqrt{1-e^2}}{na^2\left[1+\sqrt{1-e^2}\right]}\frac{\partial R}{\partial \epsilon} - \frac{h\tan(i/2)}{na^2\sqrt{1-e^2}}\frac{\partial R}{\partial i} \qquad (20\text{-}44)$$

Note that

$$1 - e^2 = 1 - h^2 - k^2 \tag{20-45}$$

These equations are, of course, good for large eccentricities as well.

For small inclinations (near equatorial orbits) we use the new variables

$$\begin{aligned} p &= \tan i \sin \Omega \\ q &= \tan i \cos \Omega \end{aligned} \tag{20-46}$$

The equations for $\frac{di}{dt}$ and $\frac{d\Omega}{dt}$ are then replaced by the following:

$$\frac{dp}{dt} = \frac{1}{na^2\sqrt{1-e^2}\cos^3 i}\frac{\partial R}{\partial q} - \frac{p}{2na^2\sqrt{1-e^2}\cos i\cos^2\frac{i}{2}}\left(\frac{\partial R}{\partial \epsilon} + \frac{\partial R}{\partial \tilde{\omega}}\right) \tag{20-47}$$

$$\frac{dq}{dt} = -\frac{1}{na^2\sqrt{1-e^2}}\frac{1}{\cos^3 i}\frac{\partial R}{\partial p} - \frac{q}{2na^2\sqrt{1-e^2}\cos i\cos^2 i/2}\left(\frac{\partial R}{\partial \epsilon} + \frac{\partial R}{\partial \tilde{\omega}}\right) \tag{20-48}$$

$$\frac{1}{\cos i} = \sqrt{1+p^2+q^2} \qquad \frac{1}{2\cos^2\frac{i}{2}} = \frac{1}{1+(1+p^2+q^2)^{-1/2}} \tag{20-49}$$

These latter expressions for the cos i can be expanded in powers of $(p^2 + q^2)$ when i is sufficiently small.

We also have problems with polar orbits where $i \approx \frac{\pi}{2}$. For these cases one usually changes the orientation of the xyz axis system.

For both near circular <u>and</u> near equatorial orbits one can use the elements,

$(a, e \sin(\Omega+\omega), e \cos(\Omega+\omega), \tan i \sin \Omega, \tan i \cos \Omega, \epsilon)$

Another set of parameters which have no singular orbits are those formed by using quaternions as elements. See, "A Nonsingular Set of Orbit Elements" by C. J. Cohen and E. C. Hubbard, U. S. Naval Weapons Laboratory, Dahlgren, Virginia, July 1961 (AD 267155).

The trouble with these nonsingular elements is that one loses physical significance; they are not as intuitively helpful as the elliptical elements. One could just as well return to canonical variables.

K. Stumpff[6], Samuel Herrick[3] and Samuel Pines[4] have each developed a variation of parameters set which use initial position and velocity vectors in the osculating plane as the state vectors. The positions are related to the instantaneous position and velocity through Kepler's equation. Using θ as the difference between the instantaneous and initial eccentric anomaly, Kepler's equation can be written,

$$nt = \theta - \left(1 - \frac{r_o}{a}\right) \sin\theta + (1 - \cos\theta) \frac{\bar{R}_o \cdot \dot{\bar{R}}_o}{a^2 n}$$

where

$\bar{R}_o$ = initial position vector

$\dot{\bar{R}}_o$ = initial velocity vector

r_o = scalar magnitude of $\bar{R}_o$.

$$\theta = E - E_o$$

The technique makes use of the fact that in unperturbed motion one can express the conic motion in the form

$$\bar{R} = f\,\bar{R}_o + g\,\dot{\bar{R}}_o$$

$$\dot{\bar{R}} = \dot{f}\,\bar{R}_o + \dot{g}\,\dot{\bar{R}}_o$$

where f and g, $\dot{f}$ and $\dot{g}$ are given in terms of initial conditions and increment of time from initial time, $t - \tau_o$, as follows

$$f = \frac{a}{r_o}(\cos\theta - 1) + 1$$

$$g = t - \tau_o - \sqrt{\frac{a^3}{\mu}}\,(\theta - \sin\theta)$$

$$\dot{f} = -\frac{\sqrt{\mu a}}{r r_o}\sin\theta$$

$$\dot{g} = \frac{a}{r}(\cos\theta - 1) + 1\ .$$

For details one should see the article by Pines in the February 1961 issue of Astronomical Journal.[4] This particular set of variables does have the difficulty that there is a deterioration in the conditioning of the associated transition matrices. This latter difficulty was overcome in a subsequent report by Pines, Wolf and Bailie[5] in which they chose a slightly different but related set of parameters. See the report referenced below.

REFERENCES

1. Cohen, C.J. and E.C. Hubbard, "A Nonsingular Set of Orbit Elements", U.S. Naval Weapons Laboratory Report (July 1961) (AD 267155).

2. Brouwer, D. and C.M. Clemence, "Methods of Celestial Mechanics", Academic Press (1961).

3. Herrick, S., Astrodynamics, D. Van Nostrand Co, Inc, (In Press).

4. Pines, S., "Variation of Parameters for Elliptic and Near Circular Orbits" Astronomical Journal, Vol 66 (February 1961) p 5-7.

5. Pines, S., H. Wolf, A. Bailie, "A Set of Uniform Variational Parameters for Space Trajectory Analysis" Analytical Mechanics Associates, Inc., Report No. 67-5 (February 1967).

6. Stumpff, K., "Himmelsmechanik" VEB Verlag, Berlin (1959).

21. OBLATE EARTH POTENTIAL

Let U be the potential corresponding to a force field $F = \nabla U$. For a solid made up of many different elemental pieces, this potential is expressed as

$$U(x,y,z) = \iiint_V \frac{dm}{R} = \iiint \frac{\sigma \, d\xi \, d\eta \, d\zeta}{\sqrt{(\xi-x)^2+(\eta-y)^2+(\zeta-z)^2}} \qquad (21\text{-}1)$$

$\sigma = \sigma(\xi,\eta,\zeta)$ is the variable density

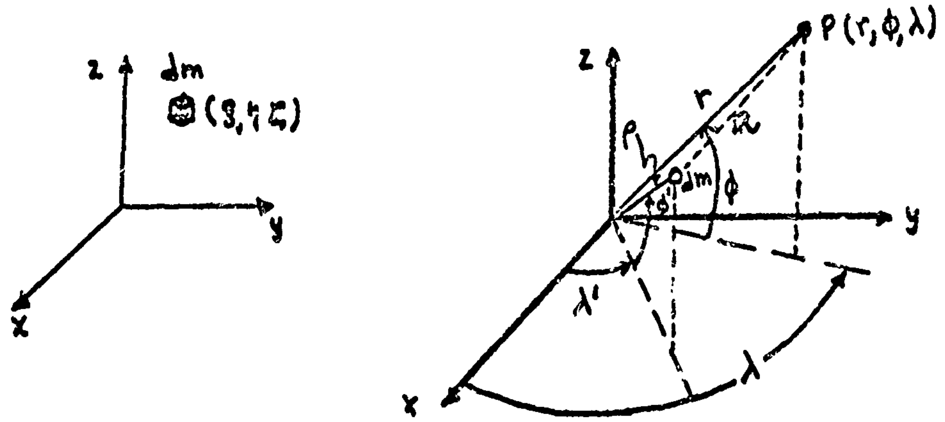


FIGURE 21-1

In spherical coordinates the unit volume is

$\rho \cos \phi' \, d\phi' \, d\lambda' \, d\rho$

$$U(r,\phi,\lambda) = \iiint \frac{\sigma \, dm}{R} = \iiint \frac{\sigma \rho^2 \cos\phi' \, d\phi' \, d\lambda' \, d\rho}{R} \qquad (21\text{-}2)$$

To solve we expand $\frac{1}{R}$

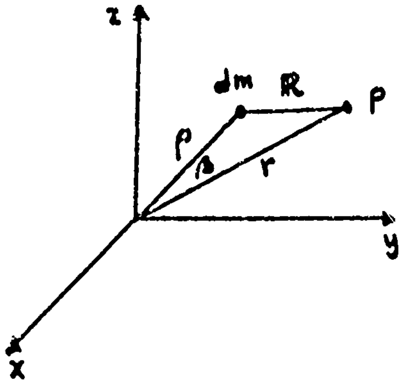


FIGURE 21-2

$$\frac{1}{R} = \frac{1}{\sqrt{\rho^2 + r^2 - 2r\rho\cos\beta}} = \frac{1}{r}\,\frac{1}{\sqrt{1 + \left(\frac{\rho}{r}\right)^2 - 2\left(\frac{\rho}{r}\right)\cos\beta}} \tag{21-3}$$

$$\frac{1}{R} = \frac{1}{r}\left\{ 1 + \frac{\rho}{r}\cos\beta + \left(\frac{\rho}{r}\right)^2\left[-\frac{1}{2} + \frac{3}{2}\cos^2\beta\right] + \right. \tag{21-4}$$

$$\left. + \left(\frac{\rho}{r}\right)^3\left[-\frac{3}{2}\cos\beta + \frac{5}{2}\cos^3\beta\right] + \left(\frac{\rho}{r}\right)^5\left[\frac{3}{8} - \frac{15}{4}\cos^2\beta + \frac{35}{8}\cos^4\beta\right] + \cdots \right\}$$

The trigonometric series in each bracket can be identified as Legendre polynomials of the first kind $P_n(\cos\beta)$. See for example, R. V. Churchill, "Fourier Series and Boundary Value Problems," Chapter 9.

Compare the list below with the coefficients of each $\left(\frac{\rho}{r}\right)^n$ term of equation (21-4).

$P_o(\cos\beta) = 1 \quad P_1(\cos\beta) = \cos\beta \quad P_2(\cos\beta) = \frac{1}{2}[3\cos^2\beta - 1]$ (21-5)

$P_3(\cos\beta) = \frac{1}{2}[5\cos^3\beta - 3\cos\beta] \quad P_4(\cos\beta) = \frac{1}{8}[35\cos^4\beta - 30\cos^2\beta + 3]$

etc.

Other terms are generated by

$$nP_n(\cos\beta) = (2n-1)\cos\beta\, P_{n-1}(\cos\beta) - (n-1)P_{n-2}(\cos\beta), \quad (21\text{-}6)$$

Thus we can write the $\frac{1}{R}$ term as

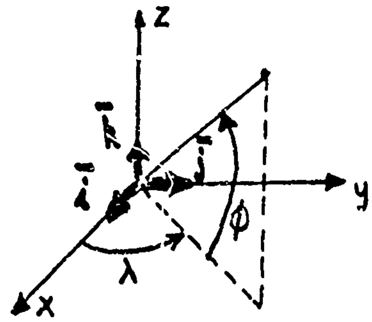
$$\frac{1}{R} = \frac{1}{r}\left[1 + \left(\frac{\rho}{r}\right)P_1(\cos\beta) + \left(\frac{\rho}{r}\right)^2 P_2(\cos\beta) + \left(\frac{\rho}{r}\right)^3 P_3(\cos\beta) + \cdots\cdots\right] \quad (21\text{-}7)$$

$$\frac{1}{R} = \frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{\rho}{r}\right)^n P_n(\cos\beta). \quad (21\text{-}8)$$

Now substitute for $\frac{1}{R}$ in the integral for U.

$$U = \frac{\sigma}{r}\sum_{n=0}^{\infty}\iiint\left(\frac{\rho}{r}\right)^n \rho^2 P_n(\cos\beta)\cos\phi'\,d\phi'\,d\lambda'\,d\rho \quad (21\text{-}9)$$

where (ρ, ϕ', λ') are the coordinates of the differential mass element dm. We now need to express $\cos\beta$ in terms of ϕ and λ, the angles of test mass at $P(r, \phi, \lambda)$.



$$\bar{r} = r\,[\cos\phi \cos\lambda\ \bar{i} + \cos\phi \sin\lambda\ \bar{j} + \sin\phi\ \bar{k}] \qquad (21\text{-}10)$$

$$\bar{\rho} = \rho\,[\cos\phi' \cos\lambda'\ \bar{i} + \cos\phi' \sin\lambda'\ \bar{j} + \sin\phi'\ \bar{k}] \qquad (21\text{-}11)$$

and hence

$$\cos\beta = \frac{\bar{r}\cdot\bar{\rho}}{r\rho} = \cos(\lambda - \lambda')\cos\phi\cos\phi' + \sin\phi\sin\phi' \qquad (21\text{-}12)$$

We now have enough information to proceed by putting (21-12) into each term of (21-9) in each of the $P_n(\cos\beta)$ expressions and then integrate term by term with respect to $d\phi'$, $d\lambda'$ and $d\rho$. (Good Grief!)

Fortunately all this is not necessary. We can make use of potential theory to solve equation (21-1)

From equation (21-1) one can find the following:

$$\frac{\partial U}{\partial x} = \iiint \frac{\sigma(\xi - x)\,d\xi\,d\eta\,d\zeta}{R^3}$$

and

$$\frac{\partial^2 U}{\partial x^2} = \iiint \sigma\left(\frac{3(\xi - x)^2}{R^5} - \frac{1}{R^3}\right) d\xi\,d\eta\,d\zeta.$$

One can find similar expressions for $\frac{\partial^2 U}{\partial y^2}$ and $\frac{\partial^2 U}{\partial z^2}$ and by addition we have

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0 \qquad (21\text{-}13)$$

which is Laplace's equation. The potential function of a distributed mass satisfies Laplace's equation in the space external to the mass. In spherical polar coordinates, (r,λ,ϕ), Laplace's equation becomes

$$\frac{\partial}{\partial r}\left(r^2 \frac{\partial U}{\partial r}\right) + \frac{1}{\cos\phi}\frac{\partial}{\partial \phi}\left(\cos\phi \frac{\partial U}{\partial \phi}\right) + \frac{1}{\cos^2\phi}\frac{\partial^2 U}{\partial \lambda^2} = 0. \quad (21\text{-}14)$$

To solve we use the separation of variables method and assume

$$U(r,\lambda,\varphi) = R(r)\Lambda(\lambda)\Phi(\phi) \quad (21\text{-}15)$$

When we substitute this into (21-14) it becomes

$$\frac{d}{dr}\left(r^2 \frac{dR}{dr}\right)\Lambda\,\Phi + \frac{1}{\cos\phi}\frac{d}{d\phi}\left(\cos\phi\frac{d\Phi}{d\phi}\right)R\,\Lambda + \frac{1}{\cos^2\phi}\frac{d^2\Lambda}{d\lambda^2}R\,\Phi = 0$$

and dividing this by $R\Lambda\Phi$ it reduces to

$$\frac{1}{R}\frac{d}{dr}\left(r^2\frac{dR}{dr}\right) = -\frac{1}{\Phi\cos\phi}\frac{d}{d\phi}\left(\cos\phi\frac{d\Phi}{d\phi}\right) - \frac{1}{\Lambda\cos^2\phi}\frac{d^2\Lambda}{d\lambda^2}. \quad (21\text{-}16)$$

Since the left hand side of (21-16) is a function of r alone, and the right hand side a function of ϕ and λ alone, these two sides must be equal to a constant which we arbitrarily select to be $q = n(n + 1)$. We can then write (21-16) in the equivalent form of

$$r^2\frac{d^2R}{dr^2} + 2r\frac{dR}{dr} - n(n+1)R = 0 \quad (21\text{-}17)$$

and

$$\frac{1}{\Lambda}\frac{d^2\Lambda}{d\lambda^2} = -q\cos^2\phi - \frac{\cos\phi}{\Phi}\frac{d}{d\phi}\left(\cos\phi\frac{d\Phi}{d\phi}\right). \quad (21\text{-}18)$$

Similarly since (21-18) has all the λ dependent terms on the left hand side, this equation must be equal to a constant which we can call $-p^2$. Equation (21-18) is then written as

$$\frac{d\Lambda}{d\lambda^2} + p^2\Lambda = 0 \tag{21-19}$$

$$\cos\phi \frac{d}{d\phi}\left(\cos\phi \frac{d\Phi}{d\phi}\right) + (q \cos^2\phi - p^2)\,\Phi = 0\,. \tag{21-20}$$

Equation (21-19) has a general solution of the form

$$\begin{aligned} \Lambda &= A \cos p\lambda + B \sin p\lambda \qquad \text{for } p \neq 0 \\ \Lambda &= C + D\lambda \qquad\qquad\qquad \text{for } p = 0\,. \end{aligned} \tag{21-21}$$

For our case we will have an axially symmetric body and hence the solution is independent of λ; therefore $p=0$, $D=0$ and $\Lambda=C$, a constant. Equation (21-20) then becomes

$$\cos\phi \frac{d}{d\phi}\left(\cos\phi \frac{d\Phi}{d\phi}\right) + q \cos^2\phi\,\Phi = 0. \tag{21-22}$$

Letting $\mu=\sin\phi$ this equation may be rewritten in the form known as Legendre's equation.

$$(1 - \mu^2)\frac{d^2\Phi}{d\mu^2} - 2\mu\frac{d\Phi}{d\mu} + q\,\Phi = 0. \tag{21-23}$$

A solution to (21-23) is obtained in the form of a power series in μ which we hope will be a convergent one at least for $|\mu| < 1$, e.g.,

$$\Phi = P_k = \sum_{k=o}^{\infty} a_k \mu^k \tag{21-24}$$

Substitution of (21-24) into (21-23) yields the recursive relation between coefficients,

$$a_{k+2} = \frac{k(1 + k) - q}{(k + 1)(k + 2)} a_k; \; k = 0,1,\ldots. \qquad (21\text{-}25)$$

The relation (21-25) leaves two arbitrary constants open to choice as is needed for the solution to (21-23). This solution can thus be written as

$$\Phi = a_o u_q(\mu) + a_1 v_q(\mu)$$

where $u_q(\mu)$ and $v_q(\mu)$ are series in even and odd powers of μ respectively. Let q be chosen such that n is an integer where $q = n(n + 1)$. For $k=n$ we have $q = k(k + 1)$ and the coefficient a_{k+2} from equation (21-25) becomes zero. Consequently the coefficients a_{k+4}, a_{k+6}, etc., will also be zero and depending on whether k is even or odd, the even or odd series, $u_q(\mu)$ or $v_q(\mu)$, terminates as an n^{th} order polynomial.

For any other values of q (where n is not an integer), the series does not terminate and can be shown to be divergent in the range $|\mu| \leq 1$ by means of Raabe's test. It therefore follows that the only solutions which remain convergent are those even or odd series terminating in an n^{th} order exponent of μ, i.e., the relevant solutions are polynomials in μ known as the Legendre polynomials of order n where n is an integer. These polynomials are given by equation (21-5).

The general solution of (21-17) can be written as

$$R = c_1 r^n + c_2 r^{-(n+1)} \qquad \text{for } n \neq -1/2$$
$$R = c_1 r^{-1/2} + c_2 r^{-1/2} \ln r \qquad \text{for } n = -1/2 \, . \tag{21-27}$$

Combining the solutions of R, Φ and Λ gives

$$U(r,\lambda,\phi) = c[c_1 r^n + c_2 r^{-(n+1)}] \; P_n(\mu) \tag{21-28}$$

where $P_n(\mu)$ are Legendre polynomials in μ. To have a potential function it is necessary that U vanish as r approaches infinity and hence the solution containing $c_1 r^n$ is neglected. Further since Laplace's equation is linear in U, linear combinations of (21-28) are also solutions. Thus we have

$$U = \sum_{n=o}^{\infty} \frac{\tilde{J}_n}{r^{n+1}} P_n(\sin \phi). \tag{21-29}$$

For a point mass we have the n=0 case and, of course $U = \mu/r$. To accommodate this case we can write ($J_o = 1$),

$$U = \frac{\mu}{r} \sum_{n=o}^{\infty} \frac{\hat{J}_n}{r^n} P_n \; (\sin \phi) \, . \tag{21-30}$$

We could have considered the series expansion in terms of $\frac{a_e}{r}$ where a_e is the equatorial radius of the earth and in that case the result could be put in the form

$$U = \frac{\mu}{r} \sum_{n=o}^{\infty} \hat{J}_n \left[\frac{a_e}{r}\right]^n P_n \; (\sin \phi) \tag{21-31}$$

or,

$$U = \frac{\mu}{r} [1 + \sum_{n=2}^{\infty} \hat{J}_n \left(\frac{a_e}{r}\right)^n P_n (\sin \phi)]. \qquad (21\text{-}32)$$

In (21-32) we neglect the term $P_1(\sin \phi)$ since it can be set equal to zero by proper choice of the origin of the coordinate system, i.e. when origin coincides with the center of gravity of the earth, then $P_1(\sin \phi) = 0$. See Chapter 35 for further details. For a spherical earth all the $\hat{J}_n$'s would be zero. We could calculate these $\hat{J}_n$ constants as functions of a_e and the shape and density of the earth by proceeding with the integration of equation (21-9). However, it is much more desirable to use dynamical measurements to calculate the $\hat{J}_n$ constant. See Chapter 35 for details of how this is done. When one calculates these $\hat{J}_n$ values, $\hat{J}_1$ turns out to be the important one, and it turns out to be negative. We would rather work with positive values so we define

$$J_n = -\hat{J}_n$$

and then the potential function can be written

$$U = \frac{\mu}{r} [1 - \sum_{n=2}^{\infty} J_n \left(\frac{a_e}{r}\right)^n P_n (\sin \phi)]. \qquad (21\text{-}33)$$

To compute the above we use the following

$$P_2(\sin \phi) = \frac{1}{2} (3 \sin^2 \phi - 1)$$

$$P_3(\sin \phi) = \frac{1}{2} (5 \sin^3 \phi - 3 \sin \phi)$$

$$nP_n(\sin \phi) = (2n-1)\sin\phi \, P_{n-1} (\sin \phi) - (n-1) P_{n-2} (\sin \phi)$$

22. SATELLITE PERTURBATIONS DUE TO AN OBLATE EARTH

The principal effect of the oblate earth can be found by using only the J_2 term. J_2 is of the order of 10^{-3} while the other J_n terms are of the order of 10^{-6} or smaller. We thus take the potential to be

$$U = \frac{\mu}{r}\left[1 + \frac{J_2}{2}\left(\frac{a_e}{r}\right)^2\left(1 - 3\sin^2\phi\right)\right] \tag{22-1}$$

With this potential, the disturbing function becomes

$$R = \frac{\mu J_2 a_e^2}{2r^3}\left[1 - 3\sin^2\phi\right] \tag{22-2}$$

The first step is to express R in terms of the orbital elements. Using equation (16-76) we have

$$\sin(f+\omega) = \frac{\sin\phi}{\sin i} \tag{22-3}$$

so that

$$R = \frac{\mu J_2 a_e^2}{2r^3}\left[1 - 3\sin^2(f+\omega)\sin^2 i\right]. \tag{22-4}$$

Using the identity

$$\sin^2(f+\omega) = \frac{1}{2}\left[1 - \cos 2(f+\omega)\right] = \frac{1}{2}\left[1 - \cos 2\omega\cos 2f + \sin 2\omega\sin 2f\right], \tag{22-5}$$

we can write R as

$$R = \frac{\mu J_2 a_e^2}{4a^3}\left[\,(2-3\sin^2 i)\left(\frac{a}{r}\right)^3 + 3\sin^2 i\,\cos 2\omega\left(\frac{a}{r}\right)^3\cos 2f + \right.$$

$$\left. - 3\sin^2 i\,\sin 2\omega\left(\frac{a}{r}\right)^3\sin 2f\,\right] \qquad (22\text{-}6)$$

where a is the semi-major axis of the orbit.

For the combinations $\left(\frac{a}{r}\right)^3$, $\left(\frac{a}{r}\right)^3 \cos 2f$, $\left(\frac{a}{r}\right)^3 \sin 2f$, there are ready-made power series expansions in multiples of the mean anomaly. (For example, see Design Guide to Orbital Flight, etc.)

(22-7)

$$\left(\frac{a}{r}\right)^3 = 1 + \frac{3}{2}e^2 + \frac{15}{8}e^4 + \ldots. + \left(3e + \frac{27}{8}e^3 + \ldots\right)\cos M$$
$$+ \left(\frac{9}{2}e^2 + \frac{7}{2}e^4 + \ldots\right)\cos 2M + \ldots.$$

$$\left(\frac{a}{r}\right)^3\cos 2f = \left(-\frac{1}{2}e + \frac{1}{12}e^3 + \ldots\right)\cos M + \left(1 - \frac{5}{2}e^2 + \frac{41}{48}e^4 + \cdots\right)\cos 2M$$
$$+ \left(\frac{7}{2}e - \frac{123}{16}e^3 + \ldots\right)\cos 3M + \ldots..$$

$$\left(\frac{a}{r}\right)^3\sin 2f = \left(-\frac{1}{2}e + \frac{1}{24}e^3 - \ldots\right)\sin M + \left(1 - \frac{5}{2}e^2 + \frac{37}{48}e^4 + \ldots\right)\sin 2M$$
$$+ \left(\frac{7}{2}e - \frac{123}{16}e^3 + \ldots\right)\sin 3M + \ldots.$$

Making use of these equations, we insert them into equation (22-6) and make the necessary rearrangements. The result will be of the form

$$R = \frac{\mu J_2 a_e^2}{a^3} \sum_{j,j'} F_{j,j'}(e, \sin i) \cos\left[jM + j'\omega\right] \tag{22-8}$$

where the coefficients $F_{j,j'}$ are series in powers of e, multiplied by certain finite expressions of sin i. See the article by Proskurin and Batrakov in list of references for an expression accurate to e^6. The following form of R is accurate to order of e^3, i.e., e^4 and higher powers have been neglected.

$$\begin{aligned} R = {} & \frac{\mu J_2 a_e^2}{2a^3}\left(1-\frac{3}{2}\sin^2 i\right)\Big[1+\frac{3}{2}e^2+3\left(e+\frac{9}{8}e^3\right)\cos M \\ & + \frac{9}{2}e^2\cos 2M + \frac{53}{8}e^3\cos 3M\Big] + \frac{3\mu J_2 a_e^2}{4a^3}\sin^2 i\Big[\frac{1}{48}e^3\cos(M-2\omega) \\ & - \frac{1}{2}\left(e-\frac{1}{8}e^3\right)\cos(M+2\omega) + \left(1-\frac{5}{2}e^2\right)\cos(2M+2\omega) \\ & + \frac{7}{2}\left(e-\frac{123}{56}e^3\right)\cos(3M+2\omega) + \frac{17}{2}e^3\cos(4M+2\omega) + \frac{845}{48}e^3\cos(5M+2\omega)\Big]. \end{aligned} \tag{22-9}$$

From this we then form the proper partial derivatives of R with respect to each of the orbital elements, $\frac{\partial R}{\partial e}$, etc., and substitute the resultant series expansion appropriately into the Lagrange planetary equations. We can then integrate term by term to obtain expressions

for the first order perturbation of the elliptic elements of the satellite due to an oblate earth.

This work is carried out by Proshurin and Batrakov (AD-245905) or (AD-250841) and is also available as NASA Technical Translation F-45 (Nov 1960).

As an example, $d_1\Omega$ is given below, again accurate to order of e^3.

$$
\begin{aligned}
d_1\Omega = -\frac{3J_2}{2}\left(\frac{a_e}{a}\right)^2 \cos i \Big\{ & (1+2e^2)nt + 3\left(e+\frac{13}{8}e^3\right)\sin M \\
& + \frac{9}{4}e^2 \sin 2M + \frac{53}{24}e^3 \sin 3M + \frac{1}{2}\left(e+\frac{3}{8}e^3\right)\sin(M+2\omega) \\
& - \frac{1}{48}e^3 \sin(M-2\omega) - \frac{1}{2}(1-2e^2)\sin(2M+2\omega) - \frac{7}{6}\left(e-\frac{95}{56}e^3\right)\sin(3M+2\omega) \\
& - \frac{17}{8}e^2 \sin(4M+2\omega) - \frac{169}{48}e^3 \sin(5M+2\omega) \Big\} .
\end{aligned}
\tag{22-10}
$$

Sterne points out that there is an easier way*. Let's return to the oblate earth disturbing function.

$$
R = \frac{\mu J_2 a_e^2}{2r^3}\left[1 - 3\sin^2\phi\right] . \tag{22-11}
$$

*Easy for him, difficult for you.

The partial derivatives of R with respect to the elements can be found by noticing that R is a function of r and ϕ and r, in turn, is a function of a, e, χ. ϕ is a function of i, ω, and f, which is itself a function of the elements e and χ only. By using the elliptical motion relations one can show that

$$\frac{\partial r}{\partial e} = -a\cos f \qquad \frac{\partial r}{\partial f} = \frac{a e(1-e^2)\sin f}{(1+e\cos f)^2} \tag{22-12}$$

$$\frac{\partial r}{\partial a} = \frac{r}{a} \qquad \frac{\partial f}{\partial \chi} = \frac{(1+e\cos f)^2}{\sqrt{1-e^2}}$$

$$\frac{\partial f}{\partial e} = \frac{\sin f\,(2+e\cos f)}{1-e^2} \qquad \frac{\partial \phi}{\partial f} = \frac{\sin i\,\cos(f+\omega)}{\cos\phi}$$

One can then easily obtain the following partials.

$$\frac{\partial R}{\partial a} = -\frac{3}{a}R \tag{22-13}$$

$$\frac{\partial R}{\partial e} = \frac{3\mu J_2 a_e^2}{2(1-e^2)r^3}\Big[\cos f\,(1+e\cos f)(1-3\sin^2 i\,\sin^2 u) - \sin^2 i\,\sin 2u\,\sin f\,(2+e\cos f)\Big]$$

$$\frac{\partial R}{\partial i} = -\frac{3J_2\mu a_e^2}{r^3}\sin i\cos i\,\sin^2 u$$

$$\frac{\partial R}{\partial \Omega} = 0$$

(22-13 continued)

$$\frac{\partial R}{\partial \omega} = -\frac{3\mu J_2 a_e^2}{2r^3} \sin^2 i \ \sin 2u$$

$$\frac{\partial R}{\partial \chi} = -\frac{3\mu J_2 a_e^2 (1+e\cos f)^4}{a^3 (1-e^2)^{7/2}} \left[(1+e\cos f)(\sin u \cos u \sin^2 i) + \frac{e \sin f}{2} (1 - 3\sin^2 u \sin^2 i) \right]$$

We can now substitute these as required in the Lagrange planetary equations to obtain the following set of exact equations.

$$\frac{da}{dt} = \frac{-\lambda e (1+e\cos f)^4 \sin f}{a^2 (1-e^2)^{7/2}} \left[1 - 3\sin^2 u \sin^2 i \right] - \frac{2\lambda (1+e\cos f)^5}{a^2 (1-e^2)^{7/2}} \sin u \cos u \sin^2 i$$

(22-14)

$$\frac{de}{dt} = -\frac{\lambda \sin f (1+e\cos f)^4}{2a^3 (1-e^2)^{7/2}} \left[1 - 3\sin^2 u \sin^2 i \right] + \left[\frac{\lambda (1+e\cos f)^3}{a^3 e (1-e^2)^{5/2}} \right] \left[\sin u \cos u \sin^2 i \right] \left[1 - (1-e\cos f)^2 \right]$$

$$\frac{d\Omega}{dt} = -\frac{\lambda (1+e\cos f)^3}{a^3 (1-e^2)^{7/2}} \sin^2 u \cos i$$

$$\frac{d\omega}{dt} = \frac{\lambda(1+e\cos f)^2}{2a^3 e(1-e^2)^{7/2}} \Big\{ (1-3\sin^2 u \sin^2 i)\left(\cos f + e(2+e\cos f)\right)$$

$$- \left[1-3\sin^2 u \sin i\right]\left[e\sin^2 f\,(2+e\cos f)\right] + \left[(-2\sin f)(1+e\cos f)(2+e\cos f)\right]\left[\sin u \cos u \sin^2 i\right]\Big\}$$

$$+ \frac{\lambda(1+e\cos f)^3 \sin^2 u \cos^2 i}{a^3(1-e^2)^{7/2}} \quad .$$

$$\frac{di}{dt} = -\frac{\lambda(1+e\cos f)^3 \sin u \cos u \cos i \sin i}{a^3(1-e^2)^{7/2}} \quad . \qquad (22\text{-}14)$$

$$\frac{d\chi_1}{dt} = -\frac{\lambda(1+e\cos f)^2}{2a^3 e(1-e^2)^3} \Big\{ \left[1-3\sin^2 u \sin^2 i\right]\Big[e(2+e\cos f) + \cos f$$

$$- e\sin^2 f\,(2+e\cos f) - 2e(1+e\cos f)\Big]$$

$$- (2\sin f)\left[(1+e\cos f)(2+e\cos f)\right]\left[\sin u \cos u \sin^2 i\right]\Big\}$$

$$\text{where} \quad \lambda = 3J_2 a_e^2 \sqrt{\frac{\mu}{a}} \quad \text{and} \quad u = f+\omega \; .$$

We could now integrate these numerically to obtain an "exact" solution. However, we need to have the true anomaly, f, as a function of time. We thus need an additional relationship.

The angular momentum vector of a satellite is, by the definition of the osculating orbit, in the direction perpendicular to the instantaneous orbit plane and its magnitude is given by the sum of the angular velocity of the satellite in the orbital plane and

the angular velocity of the orbital plane in the direction perpendicular to the orbit plane. The two terms are expressed mathematically as

$$r^2 \left(\frac{du}{dt} + \frac{d\Omega}{dt} \cos i\right) = h = \sqrt{\mu p} \tag{22-15}$$

where $u = f + \omega$ and $p = a(1-e^2)$ is the conical parameter. We thus have

$$\frac{du}{dt} = \frac{df}{dt} + \frac{d\omega}{dt} \tag{22-16}$$

and substituting and solving for $\frac{df}{dt}$ gives

$$\frac{df}{dt} = \frac{\sqrt{\mu p}}{r^2} - \frac{d\omega}{dt} - \frac{d\Omega}{dt} \cos i \tag{22-17}$$

We can now use the Lagrange planetary equations to give

$$\frac{df}{dt} = \frac{\sqrt{\mu p}}{r^2} - \frac{\sqrt{1-e^2}}{e\sqrt{\mu a}} \frac{\partial R}{\partial e} + \frac{\cot i}{\sqrt{\mu a(1-e^2)}} \frac{\partial R}{\partial i} - \frac{\cot i}{\sqrt{\mu a(1-e^2)}} \frac{\partial R}{\partial i} \tag{22-18}$$

$$\frac{df}{dt} = \frac{\sqrt{\mu p}}{r^2} - \frac{\sqrt{1-e^2}}{e\sqrt{\mu a}} \frac{\partial R}{\partial e} \tag{22-19}$$

which we can write in the form,

$$\frac{df}{dt} = \sqrt{\frac{\mu}{a}} \frac{(1+e\cos f)^2}{a(1-e^2)^{3/2}} - \frac{\sqrt{1-e^2}}{e\sqrt{\mu a}} \frac{\partial R}{\partial e} \tag{22-20}$$

For our oblate earth problem this reduces to

$$\frac{df}{dt} = \sqrt{\frac{\mu}{a}}\,\frac{(1+e\cos f)^2}{a(1-e^2)^{3/2}} - \frac{\lambda(1+e\cos f)^3}{2ea^3(1-e^2)^{7/2}}\Big\{\cos f\,(1+e\cos f)(1-3\sin^2 i\,\sin^2 u)$$

$$- \sin^2 i\,\sin 2u\,\sin f\,(2+e\cos f)\Big\} \qquad (22\text{-}21)$$

This equation together with equations (22-14) may then be used to obtain a numerical solution for the orbital elements as a function of time.

An alternate and approximate approach is to transform the equations (22-14), which have derivatives with respect to time, into equations which have derivatives with respect to $f^{(o)}$, the true anomaly of the unperturbed motion. It will then be possible to integrate the equations with respect to $f^{(o)}$ and obtain the first order perturbation effect of J_2 as closed expressions involving the initial values of the elements and $f^{(o)}$.

Recall that

$$M = \sqrt{\frac{\mu}{a^3}}\,(t-\tau_0) = E - e\sin E = 2\tan^{-1}\left[\sqrt{\frac{1-e}{1+e}}\,\tan\frac{f}{2}\right] - \frac{e\sqrt{1-e^2}\,\sin f}{1+e\cos f}$$

(22-22)

If we wish to change to true anomaly as the independent variable, we can write the left hand side of Lagrange's equations as

$$\frac{da}{dt} = \frac{da}{df}\,\frac{df}{dt}, \text{ etc.,} \qquad (22\text{-}23)$$

Where we now need an expression for $\dot{f}$. If we choose f as the true anomaly in some osculating orbit valid at time t_o, then it can be related to time by equation (22-22) in terms of the unperturbed elements a_o, e_o,

etc. We call this f the "unperturbed" true anomaly and designate it by $f^{(o)}$. We can write

$$\frac{df^{(o)}}{dt} = \frac{df^{(o)}}{dE^{(o)}} \frac{dE^{(o)}}{dM^{(o)}} \frac{dM^{(o)}}{dt}$$

Recall

$M = \sqrt{\frac{\mu}{a^3}}(t-\tau_o)$ so that $\frac{dM^{(o)}}{dt} = \sqrt{\frac{\mu}{a_0^3}}$ and since $M = E - e \sin E$,
we have $M^{(o)} = E^{(o)} - e_o \sin E^{(o)}$ and can find

$$\frac{dE^{(o)}}{dM^{(o)}} = \frac{1}{1 - e_o \cos E^{(o)}}$$

Now since

$$\frac{r}{a_o} = 1 - e_o \cos E^{(o)} = \frac{1 - e_o^2}{1 + e_o \cos f^{(o)}}$$

we have

$$\cos f^{(o)} = \frac{\cos E^{(o)} - e_o}{1 - e_o \cos E^{(o)}}$$

By using the derivative of the arccosine we can compute

$$\frac{df^{(o)}}{dE^{(o)}} = \frac{\sin E^{(o)} (1 - e_o^2)}{\sin f^{(o)} (1 - e_o \cos E^{(o)})^2}$$

and then using

[See second part of Equation (3-19) on page [illegible]].

$$\sin E^{(0)} = \frac{r \sin f^{(0)}}{a_0\sqrt{1-e_0^2}} = \frac{\sqrt{1-e_0^2}\,\sin f^{(0)}}{1+e_0\cos f^{(0)}}$$

one obtains

$$\frac{df^{(0)}}{dt} = \sqrt{\frac{\mu}{a_0^3(1-e_0^2)^3}}\,\left(1+e_0\cos f^{(0)}\right)^2 \tag{22-24}$$

Relation (22-24) transforms the Lagrange planetary equations into the following set:

$$\frac{da^{(1)}}{df^{(0)}} = +\frac{2a_0^2(1-e_0^2)^{3/2}}{\mu(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial \chi}$$

$$\frac{de^{(1)}}{df^{(0)}} = \frac{a_0(1-e_0^2)^{5/2}}{\mu e_0(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial \chi} - \frac{a_0(1-e_0^2)^2}{\mu e_0(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial \omega} \tag{22-25}$$

$$\frac{d\Omega^{(1)}}{df^{(0)}} = \frac{a_0(1-e_0^2)}{\mu \sin i_0(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial i}$$

$$\frac{d\omega^{(1)}}{df^{(0)}} = \frac{a_0(1-e_0^2)^2}{\mu e_0(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial e} - \frac{a_0(1-e_0^2)\cot i_0}{\mu(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial i}$$

$$\frac{di^{(1)}}{df^{(0)}} = \frac{a_0(1-e_0^2)\cot i_0}{\mu(1+e_0\cos f^{(0)})^2}\left(\frac{\partial R}{\partial \omega} - \frac{1}{\cos i_0}\frac{\partial R}{\partial \Omega}\right)$$

$$\frac{d\chi^{(1)}}{df^{(0)}} = -\frac{a_0(1-e_0^2)^{5/2}}{\mu e_0(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial e} - \frac{2a_0^2(1-e_0^2)^{3/2}}{\mu(1+e_0\cos f^{(0)})^2}\frac{\partial R}{\partial a}$$

The student should substitute the partial derivatives of equations (22-13) into the above to determine the form of these equations for the case of the oblate earth. When these equations are thus obtained, they can be integrated to give first-order perturbations in terms of the unperturbed anomaly, i.e., $a^{(1)} = a^{(1)} [f^{(0)}]$, $e^{(1)} = e^{(1)} [f^{(0)}]$, etc.

This suggests the following procedure: If t_o is the starting epoch, we evaluate $a^{(1)} (f^{(0)})$, $e^{(1)} (f^{(0)})$, etc. between the limits $f_o^{(0)}$ and $f_1^{(0)}$. The time t_1 at the upper limit follows from equation (22-22) in terms of a_o, e_o, etc. Using this t_1 and $\hat{a}_1 = a_o + a^{(1)} (f_1^{(0)})$, etc., in equation (22-22) we can find $\hat{f}_1$, the true anomaly for the new osculating orbit. Changing the notation from $\hat{a}_1$ to a_1, etc., and $\hat{f}_1$ to $f_1^{(0)}$, we can now repeat the procedure for the next integration interval. The "updating" of orbit elements in the right-hand side of equation (22-25) amounts to a partial allowance for higher-order perturbations, while the recalculation of f at the beginning of each step represents essentially a first order perturbation of the true anomaly.

Returning to the oblate earth problem, when the relations of equations (22-13) are substituted into equations (22-25), this set may be integrated in closed form. By means of the expansions

$$\sin^2(f+\omega) = \frac{1}{2} - \frac{1}{2}\cos(2f+2\omega) \tag{22-26}$$

$$(1+e\cos f)^2 = 1 + 2e\cos f + e^2\cos^2 f = \left(1+\frac{e^2}{2}\right) + 2e\cos f + \frac{e^2}{2}\cos 2f$$

$$(1+e\cos f)^3 = 1 + 3e\cos f + 3e^2\cos^2 f + e^3\cos^3 f$$

$$= \left(1+\frac{3}{2}e^2\right) + 3e\left(1+\frac{e^2}{4}\right)\cos f + \frac{3}{2}e^2\cos 2f + \frac{e^2}{4}\cos 3f$$

and the relations

$$2 \sin x \cos y = \cos (x+y) + \cos (x-y)$$

$$2 \sin x \cos y = \sin (x+y) + \sin (x-y) \qquad (22\text{-}27)$$

$$2 \cos x \sin y = \sin (x+y) - \sin (x-y)$$

the Lagrange planetary equations may be transformed so that only terms with sine or cosine of various angles occur, which are immediately integrable. Sterne on page 122 gives the following results. The constants K are to be adjusted to provide the correct initial values a_o, e_o, etc., and $f^{(0)}$ is the true anomaly in elliptical motion in which the relation, $n_o^2 a_o^3 = \mu$ holds. The $f^{(0)}$ in these equations will in general differ from the value of f used in calculating the disturbed motion, hence we cannot readily relate these results to time, but nevertheless the important effects can clearly be seen.

$$
\begin{aligned}
a = K_a + \frac{3 J_2 a_e^2}{a_0 (1-e_0^2)^3} \Big\{ & e_0 \left(1+\tfrac{1}{4}e_0^2\right)\left(1-\tfrac{3}{2}\sin^2 i_0\right) \cos f^{(0)} \qquad (22\text{-}28) \\
& + \tfrac{3}{4} e_0 \left(1+\tfrac{1}{4}e_0^2\right) \sin^2 i_0 \cos(2\omega_0 + f^{(0)}) \\
& + \tfrac{1}{16} e_0^3 \sin^2 i_0 \cos(2\omega_0 - f^{(0)}) \\
& + \tfrac{1}{2} e_0^2 \left(1-\tfrac{3}{2}\sin^2 i_0\right) \cos 2f^{(0)} + \left(\tfrac{1}{2}+\tfrac{3}{4}e_0^2\right) \sin^2 i_0 \cos(2\omega_0 + 2f^{(0)}) \\
& + \tfrac{1}{12} e_0^3 \left(1-\tfrac{3}{2}\sin^2 i_0\right) \cos 3f^{(0)} \\
& + \tfrac{3}{4} e_0 \left(1+\tfrac{1}{4}e_0^2\right) \sin^2 i_0 \cos(2\omega_0 + 3f^{(0)}) + \tfrac{3}{8} e_0^2 \sin^2 i_0 \cos(2\omega_0 + 4f^{(0)}) \\
& + \tfrac{1}{16} e_0^3 \sin^2 i_0 \cos(2\omega_0 + 5f^{(0)}) \Big\}
\end{aligned}
$$

$$e = K_e + \frac{3 J_2 a_e^2}{2 a_0^2 (1-e_0^2)^2} \Big\{ \left(1+\tfrac{1}{4} e_0^2\right)\left(1-\tfrac{3}{2}\sin^2 i_0\right)\cos f^{(0)}$$ (22-28 cont'd)

$$+ \tfrac{1}{4}\left(1+\tfrac{11}{4} e_0^2\right)\sin^2 i_0 \cos\left(2\omega_0 + f^{(0)}\right)$$

$$+ \tfrac{1}{16} e_0^2 \sin^2 i_0 \cos\left(2\omega_0 - f^{(0)}\right) + \tfrac{1}{2} e_0 \left(1-\tfrac{3}{2}\sin^2 i_0\right)\cos 2 f^{(0)}$$

$$+ \tfrac{5}{4} e_0 \sin^2 i_0 \cos\left(2\omega_0 + 2 f^{(0)}\right) + \tfrac{1}{12} e_0^2 \left(1-\tfrac{3}{2}\sin^2 i_0\right)\cos 3 f^{(0)}$$

$$+ \tfrac{1}{12}\left(7+\tfrac{17}{4} e_0^2\right)\sin^2 i_0 \cos\left(2\omega_0 + 3 f^{(0)}\right)$$

$$+ \tfrac{3}{8} e_0 \sin^2 i_0 \cos\left(2\omega_0 + 4 f^{(0)}\right) + \tfrac{1}{16} e_0^2 \sin^2 i_0 \cos\left(2\omega_0 + 5 f^{(0)}\right) \Big\}$$

$$i = K_i + \frac{3 J_2 a_e^2 \sin 2 i_0}{8 a_0^2 (1-e_0^2)^2} \left[e_0 \cos\left(2\omega_0 + f^{(0)}\right) + \cos\left(2\omega_0 + 2 f^{(0)}\right) + \tfrac{1}{3} e_0 \cos\left(2\omega_0 + 3 f^{(0)}\right)\right]$$

$$\Omega = K_\Omega - \frac{3 J_2 a_e^2 \cos i_0}{4 a_0^2 (1-e_0^2)^2} \Big[2 f^{(0)} + 2 e_0 \sin f^{(0)} - e_0 \sin\left(2\omega_0 + f^{(0)}\right)$$

$$- \sin\left(2\omega_0 + 2 f^{(0)}\right) - \tfrac{1}{3} e_0 \sin\left(2\omega_0 + 3 f^{(0)}\right) \Big]$$

$$\omega = K_\omega + \frac{3 J_2 a_e^2}{4 a_0^2 (1-e_0^2)^2} \Big\{ \left(4 - 5\sin^2 i_0\right) f^{(0)} + \frac{1}{e_0}\Big[\left(2 - 3\sin^2 i_0\right)$$

$$+ e_0^2\left(\tfrac{7}{2} - \tfrac{17}{4}\sin^2 i_0\right) + \tfrac{1}{4} e_0^2 \sin^2 i_0 \cos 2\omega_0 \Big] \sin f^{(0)}$$

$$- \frac{1}{e_0}\left[e_0^2 + \left(\tfrac{1}{2} - \tfrac{7}{4} e_0^2\right)\sin^2 i_0 \right] \sin\left(2\omega_0 + f^{(0)}\right)$$

$$+ \left(1 - \tfrac{3}{2}\sin^2 i_0\right)\sin 2 f^{(0)} - \left(1 - \tfrac{5}{2}\sin^2 i_0\right)\sin\left(2\omega_0 + 2 f^{(0)}\right)$$

$$+ \tfrac{1}{6} e_0 \left(1 - \tfrac{3}{2}\sin^2 i_0\right)\sin 3 f^{(0)}$$

$$- \frac{1}{3 e_0}\left[e_0^2 - \left(\tfrac{7}{2} + \tfrac{19}{8} e_0^2\right)\sin^2 i_0 \right] \sin\left(2\omega_0 + 3 f^{(0)}\right)$$

$$+ \tfrac{3}{4}\sin^2 i_0 \sin\left(2\omega_0 + 4 f^{(0)}\right) + \tfrac{1}{8} e_0 \sin^2 i_0 \sin\left(2\omega_0 + 5 f^{(0)}\right) \Big\}$$

(22-28 cont'd)

$$M = K_M + \bar{n}t + \frac{3J_2 a_e^2}{2a_o^2 e_o (1-e_o^2)^{3/2}} \Big\{ -\Big[(1-\tfrac{1}{4}e_o^2)(1-\tfrac{3}{2}\sin^2 i_o)$$

$$+ \tfrac{1}{8} e_o^2 \sin^2 i_o \cos 2\omega_o \Big] \sin f^{(o)}$$

$$+ \tfrac{1}{4}(1+\tfrac{3}{2}e_o^2)\sin^2 i_o \sin(2\omega_o + f^{(o)}) - \tfrac{1}{2}e_o(1-\tfrac{3}{2}\sin^2 i_o)\sin 2f^{(o)}$$

$$- \tfrac{1}{12} e_o^2 (1-\tfrac{3}{2}\sin^2 i_o)\sin 3f^{(o)} - \tfrac{1}{12}\sin^2 i_o (7-\tfrac{1}{4}e_o^2)\sin(2\omega_o + 3f^{(o)})$$

$$- \tfrac{3}{8} e_o \sin^2 i_o \sin(2\omega_o + 4f^{(o)}) - \tfrac{1}{16} e_o^2 \sin^2 i_o \sin(2\omega_o + 5f^{(o)}) \Big\}$$

where -

$$\bar{n} = n_o \left[1 + \frac{3J_2 a_e^2}{2a_o^2 (1-e_o^2)^{3/2}} \left(1 - \tfrac{3}{2}\sin^2 i_o\right) \right] \qquad (22\text{-}29)$$

$$n_o^2 a_o^3 = \mu \qquad (22\text{-}30)$$

REFERENCES

1. Sterne, T. E., "An Introduction to Celestial Mechanics," Interscience Tracts No. 9, Interscience Publishers, New York (1960).
2. Smart, W. M., "Celestial Mechanics," Longmans, Green and Company, New York (1953).
3. Krause, H. G. L., The Secular and Periodic Perturbations of the Orbit of an Artificial Earth Satellite. Translation of paper presented at VII International Astronautic Congress, Rome, Sep 1956 (AD-208900).
4. Proskurin, V. F. and Batrakov, Yu. V., "Perturbations in the Motion of Artificial Satellites caused by the Earth's Oblateness," Translated by J. W. Palmer from Byulleten Instituta Teoreticheskai Astronomii, Vol. 7 (1960), pp. 537-548, in RAE Translation 927, December 1960 (AD-250841). See also NASA Technical Translation F-45 (November 1960) (AD-245905).

23. SECULAR PERTURBATIONS

Sometimes we desire less information than the complete motion. We may like to know only secular changes, i.e., we want to know the change in say, Ω, after one revolution. Let's consider this problem.

To use Lagrange's planetary equations we first must determine quantities like $\frac{\partial R}{\partial \alpha_s}$ as a function of α_s, f and t, where α_s (s = 1,2,..6) are the orbital elements. We then have

$$\frac{d\alpha_s}{dt} = F_s(\alpha_s,f,t) . \tag{23-1}$$

Suppose at time t_o, the satellite passes perigee. At that time all the α_s elements will have certain values. The satellite next passes perigee at $t_o + P$ where P = period. We wish to know how much each orbital element has changed during the time P

$$\alpha_s(t_o+P) - \alpha_s(t_o) = \int_{t_o}^{t_o + P} F_s(\alpha_s,f,t)\, dt . \tag{23-2}$$

Up to this point we are exact. No approximations have been made. We now assume that the α_s values that appear in the kernel of equation (23-2) are substantially constant during one period. Then we can easily evaluate the integral (sometimes). We can make this assumption because of small perturbations. We actually have

$$\frac{d\alpha_s}{dt} = \varepsilon\, F_s(\alpha_s,f,t) \tag{23-3}$$

where ε is a small parameter. In the case of the oblate earth $J_2 \approx 10^{-3}$ is the small parameter. Hence the elements $\mathbf{a}_s$ have a slow rate of change, ergo, they are practically constant over a short time, which we take to be the time P.

This integration process is made easier by using the first order theory - since we have already made this assumption. Let's start with equations (22-25). For example

$$\frac{d\Omega}{df^{(0)}} = \frac{a_0(1-e_0^2)}{\mu \sin i_0 (1+e_0 \cos f^{(0)})^2} \frac{\partial R}{\partial i} \tag{23-4}$$

From equation (22-13)

$$\frac{\partial R}{\partial i} = -\frac{3J_2 \mu a_e^2 (1+e\cos f)^3}{a^3(1-e^2)^3} \sin i \cos i \sin^2(f+\omega) \tag{23-5}$$

which we also evaluate at the initial osculating values to give

$$\frac{d\Omega}{df^{(0)}} = -\frac{3J_2 a_e^2 (1+e_0 \cos f^{(0)}) \cos i_0 \sin^2(\omega_0 + f^{(0)})}{a_0^2 (1-e_0^2)^2} \tag{23-6}$$

Using the identity

$$\sin^2(f+\omega) = \frac{1}{2}(1-\cos 2(f+\omega)) = \frac{1}{2}\left[1-\cos 2\omega \cos 2f + \sin 2\omega \sin 2f\right]$$

equation (23-6) becomes

$$\frac{d\Omega}{df^{(0)}} = -\frac{3J_2 a_e^2 \cos i_0}{2a_0^2(1-e_0^2)^2}\left(1+e_0\cos f^{(0)}\right)\left[1-\cos 2f^{(0)}\cos 2\omega_0+\sin 2f^{(0)}\sin 2\omega_0\right] \tag{23-7}$$

This may now be integrated as $(\Delta\Omega = \Omega(t_0+P) - \Omega(t_0))$,

$$\Delta\Omega = -\frac{3J_2 a_e^2 \cos i_0}{2a_0^2(1-e_0^2)^2}\left[\int_{-\pi}^{\pi}(1+e_0\cos f^{(0)})\,df^{(0)} - \cos 2\omega_0\int_{-\pi}^{\pi}(1+e_0\cos f^{(0)})\cos 2f^{(0)}df^{(0)} + \sin 2\omega_0\int_{-\pi}^{\pi}(1+e_0\cos f^{(0)})\sin 2f^{(0)}\,df^{(0)}\right] \tag{23-8}$$

The last two integrals are zero. The first has the value 2π. Therefore

$$\Delta\Omega = -\frac{3J_2 a_e^2 \cos i_0}{2a_0^2(1-e_0^2)^2}\,2\pi \tag{23-9}$$

We can consider the change in Ω in one period per period, i.e., $\frac{\Delta\Omega}{P_0}$ using $P_0 = \frac{2\pi}{n_0}$. Thus (23-9) becomes

$$\frac{\Delta\Omega}{P_0} = -\frac{3n_0 J_2 a_e^2 \cos i_0}{2a_0^2(1-e_0^2)^2} \tag{23-10}$$

Thus the line of nodes is subject to a secular perturbation. The minus sign indicates that Ω moves backwards with respect to the direction of motion along the orbit. This is called retrograde motion.

In a similar manner the student should verify that

$$\frac{\Delta\omega}{P_0} = \frac{3n_0 J_2 a_e^2\left(1-\frac{5}{4}\sin^2 i_0\right)}{a_0^2(1-e_0^2)^2} \tag{23-11}$$

Thus the perigee location advances or regresses depending on the inclination of the orbit. If $\sin^2 i_o > 0.8$ ($i_o > 63.4^\circ$) the value of ω decreases with time and vica versa.

The other elements, with the exception of χ_1, do not have any secular changes.

Let's calculate the last secular change, that of χ_1, by a slightly different method. To find the secular perturbations we are really confined to first order approximations, hence the part of R responsible for secular terms is given by the secular part of R, i.e., by the constant term in the Fourier expansion of R (in terms of M), thus

$$R_{secular} = R_s = \frac{1}{2\pi}\int_0^{2\pi} R\, dM \tag{23-12}$$

For the oblate earth

$$R = \frac{\mu J_2 a_e^2}{2r^3}\left[1 - 3\sin^2(f+\omega)\sin^2 i\right] \tag{23-13}$$

Using the relations (which the student should confirm),

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{r^3}\right)_0 dM^{0} &= \frac{1}{a_0^3(1-e_0^2)^{3/2}} \\ \frac{1}{2\pi}\int_0^{2\pi}\left(\frac{1}{r^3}\right)_0 \cos 2(\omega_0+f^{(0)})\, dM^{(0)} &= 0 \end{aligned} \tag{23-14}$$

[Hint: change variable of integration to $df^{(0)}$, see (22-24)], one finds

$$R_s = \frac{\mu J_2 a_e^2}{2a_0^3(1-e_0^2)^{3/2}}\left[1 - \frac{3}{2}\sin^2 i_0\right] \tag{23-15}$$

We see immediately from the Lagrange planetary equations that a, e, and i suffer no secular perturbations since

$$\frac{\partial R_S}{\partial \Omega_o} = \frac{\partial R_S}{\partial \omega_o} = \frac{\partial R_S}{\partial X_o} = 0 \tag{23-16}$$

Now $\chi_1 = M - \int_0^t n dt$ and hence

$$\bar{n} = \frac{dM}{dt} = n + \frac{d\chi_1}{dt} \tag{23-17}$$

$$\bar{n} = \frac{dM}{dt} = n + \frac{d\chi_1}{dt} = n - \frac{1-e^2}{e\sqrt{\mu a}}\frac{\partial R}{\partial e} - 2\sqrt{\frac{a}{\mu}}\frac{\partial R}{\partial a} \tag{23-18}$$

where n must be considered a variable even in first order work. However, recall $n = \sqrt{\mu}\, a^{-3/2}$ so

$$\frac{dn}{dt} = -\frac{3\sqrt{\mu}}{2\sqrt{a^5}}\frac{da}{dt} = -\frac{3\mu}{a^4}\frac{\partial R}{\partial \chi} \tag{23-19}$$

For the secular perturbations due to the oblate earth, $\frac{\partial R_s}{\partial \chi} = 0$ and hence n is a constant, n_o.

Using the expression for R_s we find the secular perturbation from

$$\Delta M_s = n_o - \frac{1-e_o^2}{e_o\sqrt{\mu a_o}}\frac{\partial R_s}{\partial e_o} - 2\sqrt{\frac{a_o}{\mu}}\frac{\partial R_s}{\partial a_o}$$

$$\Delta M_s = n_o\left\{1 + \frac{3 J_2 a_e^2}{2 a_o^2 (1-e_o^2)^{3/2}}\left[1 - \frac{3}{2}\sin^2 i_o\right]\right\} \tag{23-20}$$

This perturbation in M can be interpreted as a perturbation of the period, i.e., we can allow for it by using a slightly different or perturbed value of n called $\bar{n}$.

This mean anomaly M is defined as $M \equiv n(t-\tau_0)$ where τ_0 is the time of perifocal passage. In the two body case $\frac{dM}{dt} = n =$ constant. The equation for $\frac{dM}{dt}$ in this first order secular theory is also a constant for the oblate earth perturbation because $\dot{a}$, $\dot{e}$ and $\frac{di}{dt}$ experience no secular variations (see equation (23-20). Hence the unperturbed mean motion (n_0) is either increased or decreased, depending on the value of i_0, by a certain constant amount. Adopting some convenient new epoch time $t = t_0$ where $M = M_0$ we can write equation (23-20) above in the form

$$M = M_0 + \bar{n}\ (t-t_0)\ , \tag{23-21}$$

To make the mean anomaly consistent we can write this as

$$M = M_0 + \left[\frac{3 J_2 a_e^2}{2 a_0 (1-e_0^2)^{3/2}}\left(1 - \frac{3}{2}\sin^2 i_0\right)\right] n_0 (t-t_0) + n_0 (t-t_0) \tag{23-32}$$

or

$$M = M_0 + \dot{M}_0\ (t-t_0) + n_0\ (t-t_0) \tag{23-23}$$

In a similar fashion we can write for the secular changes

$$\Omega = \Omega_0 + \dot{\Omega}\ (t-t_0) \tag{23-24}$$

$$\omega = \omega_o + \dot{\omega}_o (t-t_o) \tag{23-25}$$

Thus the first order J_2 secular variations can be summarized as below.

$$\bar{n} = n_o \left[1 + \frac{3 J_2 a_e^2}{2 a_o^2 (1-e_o^2)^{3/2}} \left(1 - \frac{3}{2} \sin^2 i_o\right) \right]$$

$$\bar{P} = \frac{2\pi}{\bar{n}}$$

$$M = M_o + \bar{n}(t-t_o) = E - e_o \sin E$$

$$\Omega = \Omega_o - \left(\frac{3 J_2 a_e^2 \cos i_o}{2 a_o^2 (1-e_o^2)^2} \right) \bar{n}(t-t_o)$$

$$\omega = \omega_o + \left(\frac{3 J_2 a_e^2}{a_o^2 (1-e_o^2)^2} \left(1 - \frac{5}{4} \sin^2 i_o\right) \right) \bar{n}(t-t_o)$$

(23-26)

$a = a_o$

$$\tan \frac{f}{2} = \sqrt{\frac{1+e_o}{1-e_o}} \tan \frac{E}{2}$$

$e = e_o$

$$p_o^2 = a_o^2 (1-e_o^2)^2$$

$i = i_o$

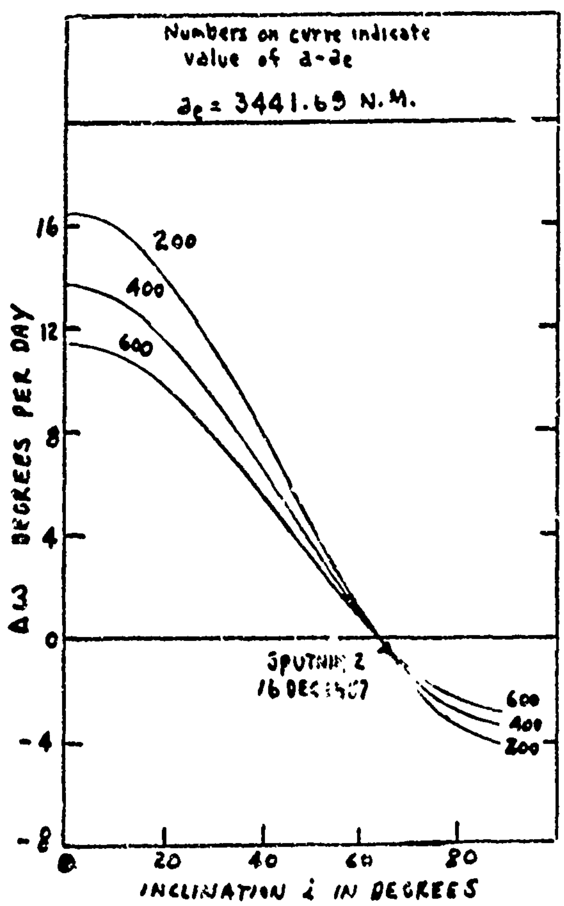
The time unit here is the unperturbed period P_0.

It is sometimes more convenient to adopt a mean value of the semi-major axis as not a_0 but

$$\bar{a} = a_0 \left[1 - \frac{3 J_2 a_e^2}{2 a_0^2 (1 - e_0^2)} \left(1 - \frac{3}{2} \sin^2 i_0 \right) \right] \tag{23-27}$$

then

$$n^2 \bar{a}^3 = \mu \left[1 - \frac{3 J_2 a_e^2}{2 a_0^2 (1 - e_0^2)} \left(1 - \frac{3}{2} \sin^2 i_0 \right) \right] \tag{23-28}$$



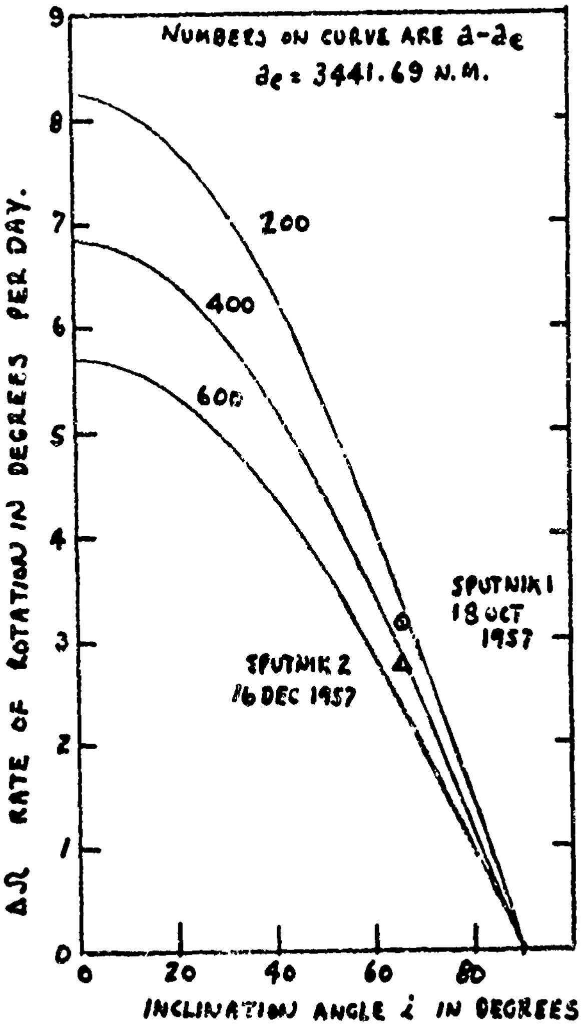
$$\frac{\Delta\omega}{P_0} = \frac{3 n_0 J_2 a_e^2 \left(1 - \frac{5}{4} \sin^2 i_0\right)}{a_0^2 (1 - e_0^2)^{3/2}}$$

to $O(e^1)$, $O(J_2)$

$$\Delta\omega = 5.0 \left(\frac{a_e}{a_0}\right)^{3.5} (5 \cos^2 i_0 - 1)$$

in degrees per day.

FIGURE 23-1. Rate of rotation of major axis of orbit, measured in the orbital plane.



$$\frac{\Delta\Omega}{P} = -\frac{3 n_0 J_2 a_e^2 \cos i_0}{2 a_0^2 (1-e_0^2)^2}$$

to $O(e^2)$, $O(J_2)$

$$\Delta\Omega = 10.0\left(\frac{a_e}{a_0}\right)^{3.5} \cos i_0$$

degrees/day.

FIGURE 23-2. Rate of rotation of the orbital plane about the Earth's axis.

24. GAUSS PLANETARY EQUATIONS

The Lagrange planetary equations can be modified in a number of ways. One of the most fruitful modifications is due to Gauss. To develop the Gauss planetary equations we first resolve the disturbing force into three components, U, V and W where W is perpendicular to the plane of the orbit, positive toward the north pole; V is in the plane of the orbit at right angles to the radius vector, positive in the direction of motion; U is along the radius vector, positive in the positive r direction. These are shown in Figure 24-1 on the next page. Note carefully that here U is a force and is not the potential function of the previous sections.

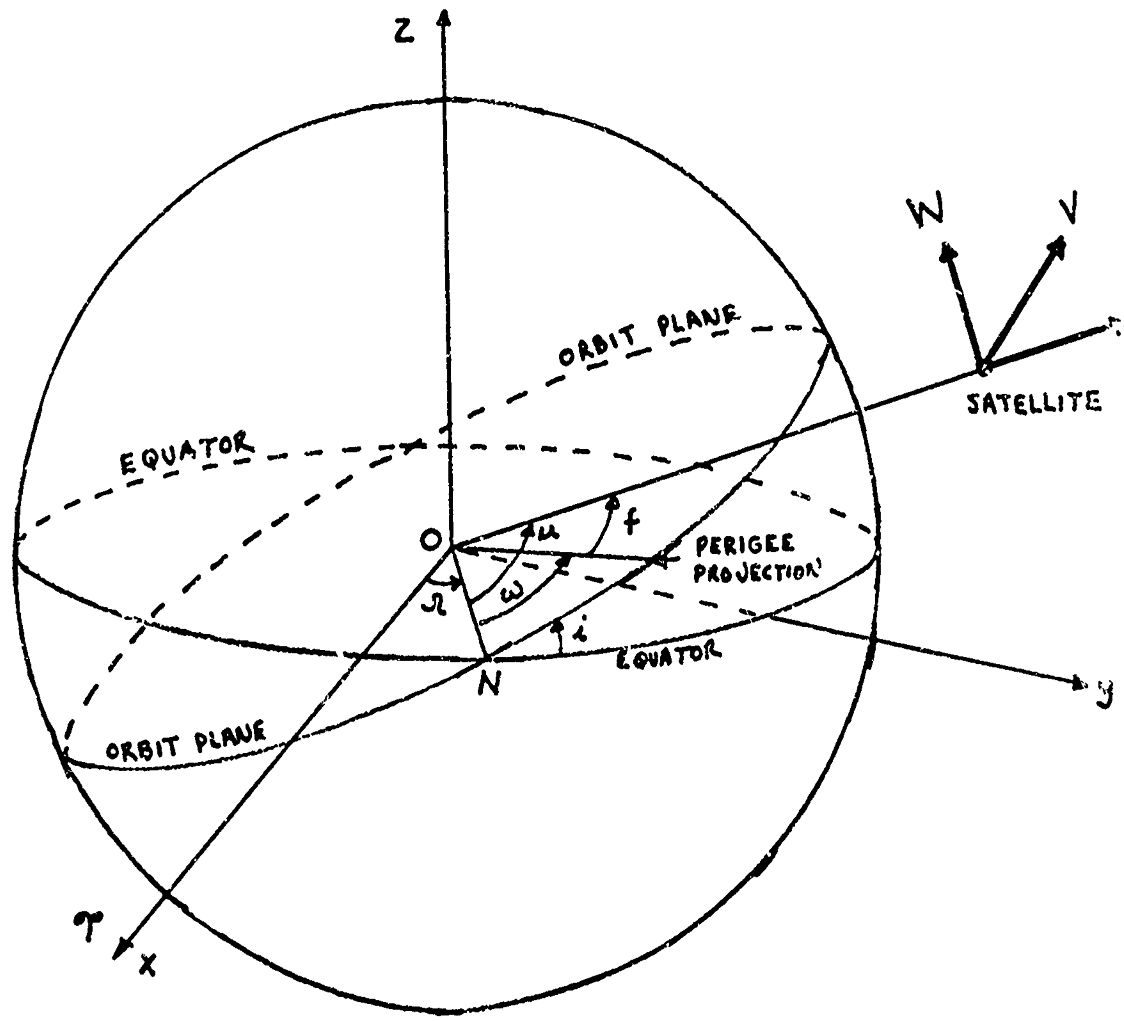
From the spherical geometry of this Figure 24-1, we can write

$$F_x = \frac{\partial R}{\partial x} = U\,[\cos u \cos \Omega - \sin u \sin \Omega \cos i] - V\,[\sin u \cos \Omega + \cos u \sin \Omega \cos i] + W \sin \Omega \sin i$$

$$F_y = \frac{\partial R}{\partial y} = U\,[\cos u \sin \Omega + \sin u \cos \Omega \cos i] - V\,[\sin u \sin \Omega - \cos u \cos \Omega \cos i] - W \cos \Omega \sin i \qquad (24\text{-}1)$$

$$F_z = \frac{\partial R}{\partial z} = U \sin u \sin i + V \cos u \sin i + W \cos i$$

These are Euler angle transformations. The $x^1y^1z^1$ system is formed from the xyz system by three rotations. (1) Rotate through $+\Omega$ about the z axis; (2) rotate through $+i$ about x^1 axis; (3) rotate through $+u = \omega + f$ about the z^1 axis. We thereby obtain



$u = \omega + f$

f = true anomaly.

FIGURE 24-1. UNIT SPHERE PROJECTION OF SATELLITE ORBIT.
W, V, U FORCES

$$
\begin{bmatrix} x^1 \\ y^1 \\ z^1 \end{bmatrix} = \begin{bmatrix} \cos u \cos \Omega - \sin u \cos i \sin \Omega & \cos u \sin \Omega + \sin u \cos i \cos \Omega & \sin i \sin u \\ -\sin u \cos \Omega - \sin \Omega \cos i \cos u & -\sin u \sin \Omega + \cos u \cos i \cos \Omega & \sin i \cos u \\ \sin i \sin \Omega & -\cos \Omega \sin i & \cos i \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{24-2}
$$

The perturbative forces in $x^1y^1z^1$ system are U, V, W

$$
\frac{\partial R}{\partial x^1} = U \; ; \qquad \frac{\partial R}{\partial y^1} = V \; ; \qquad \frac{\partial R}{\partial z^1} = W \; .
$$

Letting the elliptical elements be a_s then we can form

$$
\frac{\partial R}{\partial a_s} = \frac{\partial R}{\partial x^1}\frac{\partial x^1}{\partial a_s} + \frac{\partial R}{\partial y^1}\frac{\partial y^1}{\partial a_s} + \frac{\partial R}{\partial z^1}\frac{\partial z^1}{\partial a_s} + \frac{\partial R}{\partial \dot{x}^1}\frac{\partial \dot{x}^1}{\partial a_s} + \frac{\partial R}{\partial \dot{y}^1}\frac{\partial \dot{y}^1}{\partial a_s} + \frac{\partial R}{\partial \dot{z}^1}\frac{\partial \dot{z}^1}{\partial a_s}
$$

Since R is a function only of position, $\frac{\partial R}{\partial \dot{x}^1} = \frac{\partial R}{\partial \dot{y}^1} = \frac{\partial R}{\partial \dot{z}^1} = 0$ sooo

$$
\frac{\partial R}{\partial a_s} = U\frac{\partial x^1}{\partial a_s} + V\frac{\partial y^1}{\partial a_s} + W\frac{\partial z^1}{\partial a_s} \; . \tag{24-3}
$$

The equations (24-2) are of the form

$$
x^1 = A_x (u,\Omega,i)\, x + B_x (u,\Omega,i)\, y + C_x (u,\Omega,i)\, z
$$
$$
y^1 = A_y (u,\Omega,i)\, x + B_y (u,\Omega,i)\, y + C_y (u,\Omega,i)\, z
$$
$$
z^1 = A_z (u,\Omega,i)\, x + B_z (u,\Omega,i)\, y + C_z (u,\Omega,i)\, z
$$

From these kind of equations we first form terms of the sort

$$\frac{\partial x'}{\partial a_s} = \frac{\partial A_x}{\partial a_s} x + \frac{\partial B_x}{\partial a_s} y + \frac{\partial C_x}{\partial a_s} z + A_x \frac{\partial x}{\partial a_s} + B_x \frac{\partial y}{\partial a_s} + C_x \frac{\partial z}{\partial a_s}$$

$$\frac{\partial y'}{\partial a_s} = \frac{\partial A_y}{\partial a_s} x + \frac{\partial B_y}{\partial a_s} y + \frac{\partial C_y}{\partial a_s} z + A_y \frac{\partial x}{\partial a_s} + B_y \frac{\partial y}{\partial a_s} + C_y \frac{\partial z}{\partial a_s}$$

$$\frac{\partial z'}{\partial a_s} = \frac{\partial A_z}{\partial a_s} x + \frac{\partial B_z}{\partial a_s} y + \frac{\partial C_z}{\partial a_z} z + A_z \frac{\partial x}{\partial a_s} + B_z \frac{\partial y}{\partial a_s} + C_z \frac{\partial z}{\partial a_s} \qquad (24\text{-}4)$$

Thereupon by using (Equation (5-31) on page 50),

$$x = r \cos u \cos \Omega - r \sin u \cos i \sin \Omega$$

$$y = r \cos u \sin \Omega + r \sin u \cos i \cos \Omega \qquad (24\text{-}5)$$

$$z = r \sin u \sin i$$

$$r = \frac{a(1-e^2)}{1 + e \cos f} \qquad u = f + \omega \qquad \chi = -n\tau_0 \qquad n^2 = \frac{\mu}{a^3}$$

we finally obtain the following

$$\frac{\partial x^1}{\partial a} = \frac{r}{a} \qquad \frac{\partial y^1}{\partial a} = 0 \qquad \frac{\partial z^1}{\partial a} = 0$$

$$\frac{\partial x^1}{\partial e} = -a \cos f \qquad \frac{\partial y^1}{\partial e} = \frac{a \sin f\,(2 + e \cos f)}{1 + e \cos f} \qquad \frac{\partial z^1}{\partial e} = 0$$

$$\frac{\partial x^1}{\partial \omega} = 0 \qquad \frac{\partial y^1}{\partial \omega} = r \qquad \frac{\partial z^1}{\partial \omega} = 0 \qquad (24\text{-}6)$$

$$\frac{\partial x^1}{\partial \chi} = \frac{ea \sin f}{\sqrt{1 - e^2}} \qquad \frac{\partial y^1}{\partial \chi} = \frac{a^2\sqrt{1 - e^2}}{r} \qquad \frac{\partial z^1}{\partial \chi} = 0 \qquad \text{(24-6 contd)}$$

$$\frac{\partial x^1}{\partial i} = 0 \qquad \frac{\partial y^1}{\partial i} = 0 \qquad \frac{\partial z^1}{\partial i} = r \sin u$$

$$\frac{\partial x^1}{\partial \Omega} = 0 \qquad \frac{\partial y^1}{\partial \Omega} = r \cos i \qquad \frac{\partial z^1}{\partial \Omega} = - r \cos u \sin i$$

Combining this with (24-4) along with (24-3), we find, after much wailing and gnashing of teeth, that

$$\frac{\partial R}{\partial \Omega} = V\, r \cos i - W\, r \cos u \sin i$$

$$\frac{\partial R}{\partial i} = W\, r \sin u$$

$$\frac{\partial R}{\partial \omega} = V\, r$$

$$\frac{\partial R}{\partial a} = U \frac{r}{a} \qquad \text{(24-7)}$$

$$\frac{\partial R}{\partial e} = -U\, a \cos f + V\, [1 + \frac{r}{p}]\, a \sin f$$

$$\frac{\partial R}{\partial \chi} = \frac{Uae}{\sqrt{1 - e^2}} \sin f + V \frac{a^2}{r} \sqrt{1 - e^2}$$

$$\frac{r}{p} = \frac{1}{1 + e \cos f} \;\; ; \;\; \frac{r}{a} = \frac{1 - e^2}{1 + e \cos f} \;\; ; \;\; p = a\,(1-e^2)$$

With these equations, the Lagrange planetary equations become,

$$\frac{d\Omega}{dt} = \frac{W \sin u \sqrt{1-e^2}}{n a \sin i \, (1+e\cos f)}$$

$$\frac{di}{dt} = \frac{W \cos u \sqrt{1-e^2}}{n a \, (1+e\cos f)}$$

$$\frac{d\omega}{dt} = -\frac{U \cos f \sqrt{1-e^2}}{nae} + \frac{\sqrt{1-e^2}\,(2+e\cos f)\sin f}{nae(1+e\cos f)} V - W \frac{\sin u \cot i \sqrt{1-e^2}}{na\,(1+e\cos f)}$$

$$\frac{de}{dt} = \frac{\sqrt{1-e^2}\,\sin f}{na} U + \frac{\sqrt{1-e^2}}{nae}\left[1+e\cos f - \frac{1-e^2}{1+e\cos f}\right] V \qquad (24\text{-}8)$$

$$\frac{da}{dt} = \frac{2e\sin f}{n\sqrt{1-e^2}} U + \frac{2(1+e\cos f)}{n\sqrt{1-e^2}} V$$

$$\frac{d\chi}{dt} = \frac{1}{na}\left[\frac{2(1-e^2)}{1+e\cos f} - \frac{1-e^2}{e}\cos f\right] U$$

$$-\left\{\frac{1-e^2}{nae}\left(\frac{2+e\cos f}{1+e\cos f}\right)\sin f\right\} V$$

$$\frac{2+e\cos f}{1+e\cos f} = 1 + \frac{r}{p}$$

These (24-8) are called the Gauss planetary equations.

Another representation is in terms of T, N and W. T is in the direction of motion, along the velocity vector; N is normal to T, positive toward the interior of the ellipse; and W is normal to the orbital plane as before.

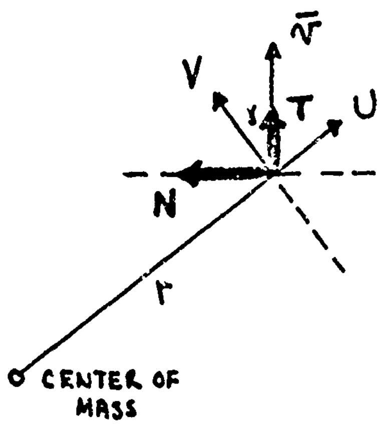


FIGURE 24-2
(In the planes of the orbit)

Recall equation (5-6),

$$\tan \gamma = \frac{e \sin f}{1 + e \cos f} \tag{24-9}$$

Where γ is the flight path angle, here represented by the angle between T and V as shown in Figure 24-2. $\bar{V}$ is the velocity vector of the particle. From the above we have

$$\sin \gamma = \frac{e \sin f}{\sqrt{1 + e^2 + 2e \cos f}}$$

(24-10)

$$\cos \gamma = \frac{1 + e \cos f}{\sqrt{1 + e^2 + 2e \cos f}}$$

and since

$$V = T\cos\gamma + N\sin\gamma$$
$$U = T\sin\gamma - N\cos\gamma \qquad (24\text{-}11)$$

we have

$$V = \frac{1+e\cos f}{\sqrt{1+e^2+2e\cos f}}\,T + \frac{e\sin f}{\sqrt{1+e^2+2e\cos f}}\,N$$

$$U = \frac{e\sin f}{\sqrt{1+e^2+2e\cos f}}\,T - \frac{1+e\cos f}{\sqrt{1+e^2+2e\cos f}}\,N \qquad (24\text{-}12)$$

Making this substitution for V and U in Gauss's equation (24-8) gives

$$\frac{d\Omega}{dt} = \frac{\sin u\sqrt{1-e^2}}{na\sin i\,(1+e\cos f)}\,W$$

$$\frac{di}{dt} = \frac{\sqrt{1-e^2}\,\cos u}{na\,(1+e\cos f)}\,W$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{nae\sqrt{1+2e\cos f+e^2}}\left[2T\sin f + N\left(2e + \frac{(1-e^2)\cos f}{1+e\cos f}\right)\right] - W\frac{(\sin u)\sqrt{1-e^2}}{na\,(1+e\cos f)\tan i}$$

$$\frac{de}{dt} = \frac{\sqrt{1-e^2}}{na\sqrt{1+2e\cos f+e^2}}\left[2T(\cos f + e) - N\frac{(1-e^2)\sin f}{1+e\cos f}\right] \qquad (24\text{-}13)$$

$$\frac{da}{dt} = \frac{2T\sqrt{1+2e\cos f+e^2}}{n\sqrt{1-e^2}}$$

$$\frac{d\chi}{dt} = \frac{1-e^2}{nae\sqrt{1+2e\cos f+e^2}}\left[2T\sin f\left(1+\frac{e^2}{1+e\cos f}\right) + N\frac{(1-e^2)\cos f}{1+e\cos f}\right]$$

Both equations (24-8) and (24-13) are called Gauss's planetary equations. As before we can change the independent variable to $f^{(o)}$ or $E^{(o)}$. The student should carry out this transformation.

25. ORBITS WITH SMALL CONSTANT THRUST

As an application of Gauss's equations consider the effect of small constant thrust which is so directed that

$$U = 0 \quad , \quad V = 0 \quad , \quad W = C \, . \tag{25-1}$$

This is an ion engine (say) which is always directed perpendicular to the orbital plane. Substituting these thrust values into equations (24-8) we see immediately that

$$\dot{a} = 0 \qquad \dot{e} = 0 \qquad \dot{\chi}_1 = 0 \, . \tag{25-2}$$

Thus the energy and shape of the orbit are not affected by this kind of perturbation. We expect this physically since the force is always perpendicular to the direction of motion and therefore cannot change the particles energy. Ω, ω and i are subject to perturbations and we wish to calculate their secular accelerations. To do this we assume C is small.

Consider first the secular change in Ω. To facilitate the calculation we change the independent variable to $f^{(o)}$ and the corresponding equation (24-8) becomes

$$\frac{d\Omega}{df^{(o)}} = \frac{C a_o^2 (1-e_o^2)^2 \sin(f^{(o)} + \omega_o)}{\mu \sin i_o (1 + e_o \cos f^{(o)})^3} \tag{25-3}$$

To find the secular change we integrate over one revolution.

$$\Delta\Omega = \Omega(t_0+P) - \Omega(t_0) = \frac{Ca_0^2(1-e_0^2)}{\mu \sin i_0}\int_{-\pi}^{\pi}\frac{\sin(f^{(0)}+\omega_0)}{(1+e_0\cos f^{(0)})^3}df^{(0)} \quad (25\text{-}4)$$

Expanding the integral

$$\Delta\Omega = \frac{Ca_0^2(1-e_0^2)}{\mu \sin i_0}\left[\sin\omega_0\int_{-\pi}^{\pi}\frac{\cos f^{(0)}}{(1+e_0\cos f^{(0)})^3}df^{(0)} + \cos\omega_0\int_{-\pi}^{\pi}\frac{\sin f^{(0)}\, df^{(0)}}{(1+e_0\cos f^{(0)})^3}\right. \quad (25\text{-}5)$$

The second integral vanishes because of odd symmetry of the integral. This leaves

$$\Delta\Omega = \frac{Ca_0^2(1-e_0^2)}{\mu \sin i_0} I(e_0) \sin\omega_0 \quad (25\text{-}6)$$

where

$$I(e_0) = \int_{-\pi}^{\pi}\frac{\cos f^{(0)}\, df^{(0)}}{(1+e_0\cos f^{(0)})^3} = \frac{3\pi e_0}{(1-e_0^2)^{5/2}} \quad (25\text{-}7)$$

The integral is a little messy to evaluate and it is at this point in any general problem of this type that one may have trouble. However, the integral could be evaluated numerically and anyway its exact value is not essential to our solution. A brief analysis shows that the I(e) function must have the general character indicated in Figure 25-1. In particular I(o) = 0.

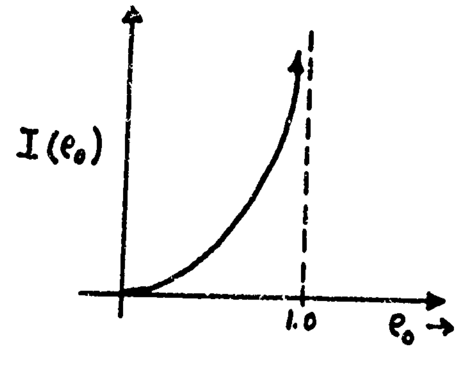


FIGURE 25-1

Applying the same process to ω and i equations gives

$$\Delta\omega = \frac{C a_o^2 (1-e_o^2)^2 I(e_o) \sin\omega_o}{\mu \tan i_o} \tag{25-8}$$

$$\Delta i = \frac{C a_o^2 (1-e_o^2)^2 I(e_o) \cos\omega_o}{\mu} \tag{25-9}$$

Note that for a circular orbit $e = 0$, $I(e) = 0$ and all the secular perturbations vanish. Does this seem reasonable?

There is nothing in the problem that dictates the orientation of the xyz inertial coordinate system. Accordingly without loss of generality we may require that the xy plane initially contain the line of apsides ($\omega_o = 0$), but not be the orbit plane itself ($i_o \neq 0$). Then

$$\Delta\Omega = 0 \qquad \Delta\omega = 0 \qquad \Delta i = \frac{C a_o^2 (1-e_o)^2 I(e_o)}{\mu} \tag{25-10}$$

This is uniform rotation of the orbit plane around the line of apsides. The rate of rotation increases with e_o, and is zero for $e_o = 0$.

By this method we are able to determine rather easily the behavior of the orbit, and in terms which are easily understood. Now the student should try the other orthogonal thrust cases, i.e.,

(1) $U = 0 \quad V = c \quad W = 0$ azimuthal thrust.

(2) $U = c \quad V = 0 \quad W = 0$ radial thrust.

Assume the thrust C is very small and find the secular changes in a, e and determine which, if any, of the remaining elements are free of secular perturbations. Which method of thrusting would allow the particle to escape the planet in the shortest time? Do the results of the analysis agree with your physical ideas of how this thrust should perturb the orbit? Explain.

REFERENCES

1. Lass, H. and C. B. Solloway, "Motion of a Satellite Under the Influence of a Constant Normal Thrust," ARS Journal, Vol 32, No. 1, page 97, January 1962.

2. Citron, S. J., "The Solutions for Satellite Motion under Low Acceleration....," ARS Journal, Vol 31, No. 12, pages 1786-1787, December 1961.

3. Tsien, H. S., "Take off from Satellite Orbits," Jet Propulsion, Vol. 23, pages 233-236, July-August 1953.

4. Dobrowolski, A., "Satellite Orbit Perturbations under a Continuous Radial Thrust of Small Magnitude," Jet Propulsion, Vol 28, pages 687-688, October 1958.

5. Lass, H. and J. Torell, "Low Acceleration Takeoff from a Satellite Orbit," ARS Journal, Vol 31, pages 24-28, January 1961.

6. Ehricke, K. A., Space Flight Vol II, Dynamics, D. van Nostrand Co. 1962.

7. Koelle, H., editor, Handbook of Astronautical Engineering, Chapter on Low Thrust by E. Levin.

26. LUNI-SOLAR PERTURBATIONS

Consider the coordinate system shown below:

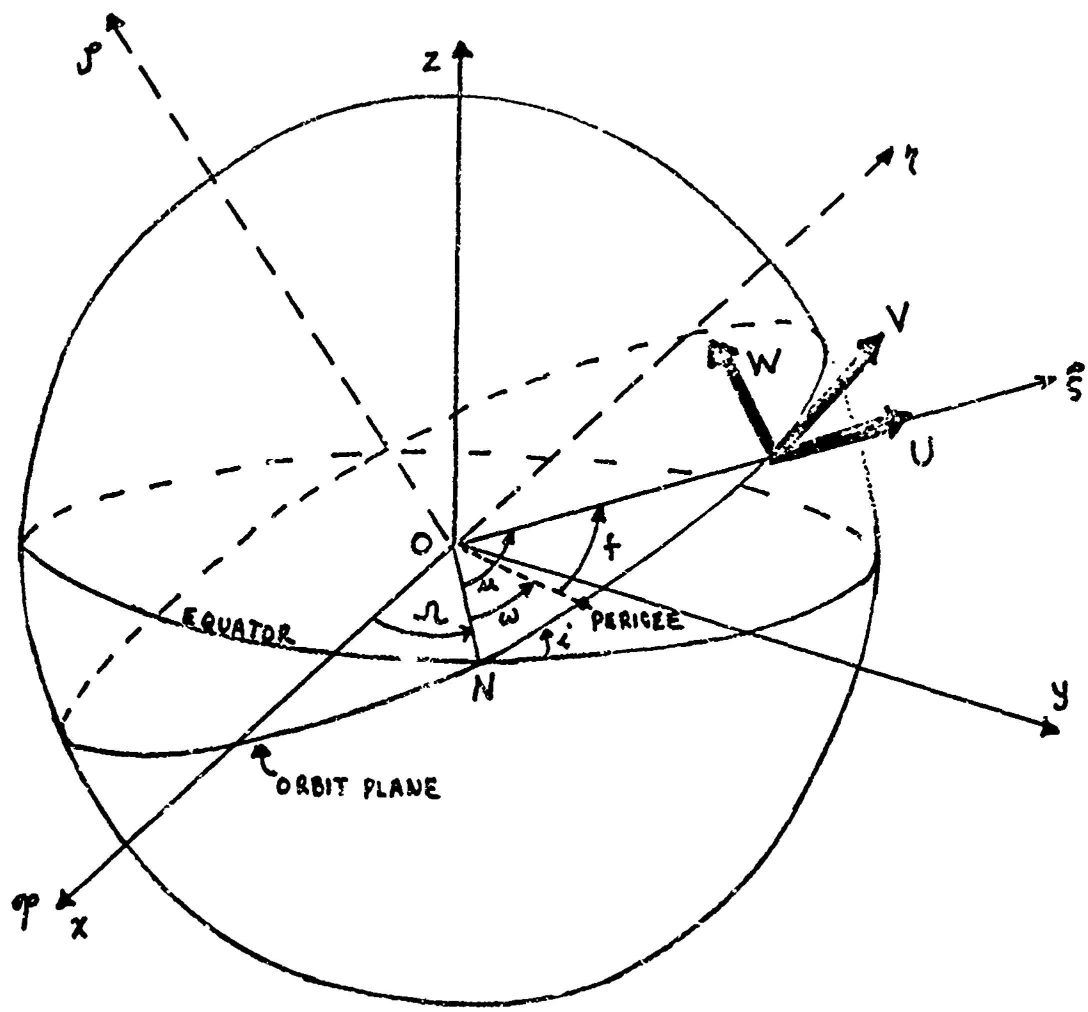


FIGURE 26-1. SATELLITE ORBIT PROJECTION ON UNIT SPHERE

Let the coordinate of the satellite be (x,y,z) and those of the disturbing body (x_d, y_d, z_d). Then the disturbing function R, whose

derivative in any direction gives the component of the disturbing <u>force</u> in that direction is

$$R = k^2 M_d \left(\frac{1}{\Delta} - \frac{x x_d + y y_d + z z_d}{r_d^3} \right) \tag{26-1}$$

where M_d is the mass of the disturbing body, which is at a distance r_d from the earth's center. Δ is the distance between satellite and disturbing body, i.e.,

$$\Delta^2 = (x - x_d)^2 + (y - y_d)^2 + (z - z_d)^2 . \tag{26-2}$$

The components of the disturbing force along the inertial axis, Ox, Oy and Oz are

$$\begin{aligned} \frac{\partial R}{\partial x} &= -k^2 M_d \left(\frac{x - x_d}{\Delta^3} + \frac{x_d}{r_d^3} \right) \\ \frac{\partial R}{\partial y} &= -k^2 M_d \left(\frac{y - y_d}{\Delta^3} + \frac{y_d}{r_d^3} \right) \\ \frac{\partial R}{\partial z} &= -k^2 M_d \left(\frac{z - z_d}{\Delta^3} + \frac{z_d}{r_d^3} \right) \end{aligned} \tag{26-3}$$

If we consider the direction cosines which denote the force axis U, V, W by (l_1, m_1, n_1), (l_2, m_2, n_2) and (l_3, m_3, n_3) respectively,

we can
write these in terms of the satellite coordinates as follows (see Section 5 of these notes on Smart, "Celestial Mechanics," page 63), (also page 235 of notes),

$$
\begin{aligned}
\ell_1 &= \cos \Omega \cos u - \sin \Omega \sin u \cos i \\
m_1 &= \sin \Omega \cos u + \cos \Omega \sin u \cos i \\
n_1 &= \sin u \sin i
\end{aligned}
\tag{26-4}
$$

$$
\begin{aligned}
\ell_2 &= -\cos \Omega \sin u - \sin \Omega \cos u \cos i \\
m_2 &= -\sin \Omega \sin u + \cos \Omega \cos u \cos i \\
n_2 &= \cos u \sin i
\end{aligned}
\tag{26-5}
$$

$$
\begin{aligned}
\ell_3 &= \sin \Omega \sin i \\
m_3 &= -\cos \Omega \sin i \\
n_3 &= \cos i
\end{aligned}
\tag{26-6}
$$

If the orbital elements of the disturbing body are denoted by subscript d, its coordinates are

$$
\begin{aligned}
x_d &= r_d (\cos \Omega_d \cos u_d - \sin \Omega_d \sin u_d \cos i_d) \\
y_d &= r_d (\sin \Omega_d \cos u_d + \cos \Omega_d \sin u_d \cos i_d) \\
z_d &= r_d (\sin u_d \sin i_d) \; .
\end{aligned}
\tag{26-7}
$$

Denoting the angle between the radius vector to the satellite and disturbing body by β, we can write

$$\Delta^2 = r^2 + r_d^2 - 2rr_d \cos \beta \qquad (26\text{-}8)$$

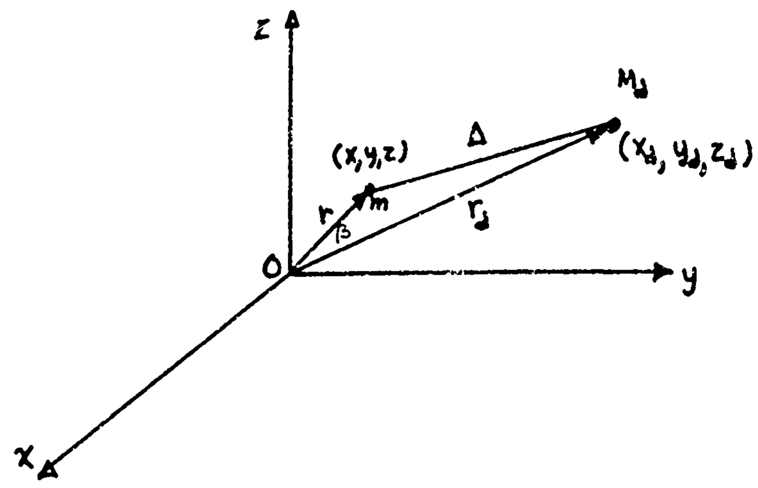


FIGURE 26-2

so that

$$\frac{1}{\Delta^3} = \frac{1}{r_d^3}\left[1 + 3\frac{r}{r_d}\cos\beta + O\left(\left(\frac{r}{r_d}\right)^2\right)\right] \qquad (26\text{-}9)$$

Using (26-4) and (26-7) we obtain

$$\cos\beta = \frac{xx_d + yy_d + zz_d}{rr_d} = A\cos u + B\sin u \qquad (26\text{-}10)$$

where

$$A = \cos(\Omega-\Omega_d)\cos u_d + \cos i_d \sin u_d \sin(\Omega-\Omega_d)\,.$$

$$(26\text{-}11)$$

$$B = \cos i\,[\cos i_d \sin u_d \cos(\Omega-\Omega_d) - \sin(\Omega-\Omega_d)\cos u_d]$$
$$+ \sin i \sin i_d \sin u_d\,.$$

Substitution of equation (26-9) into (26-3) using the relations

$$U = \ell_1 \frac{\partial R}{\partial x} + m_1 \frac{\partial R}{\partial y} + n_1 \frac{\partial R}{\partial z}$$

$$V = \ell_2 \frac{\partial R}{\partial x} + m_2 \frac{\partial R}{\partial y} + n_2 \frac{\partial R}{\partial z} \qquad (26\text{-}12)$$

$$W = \ell_3 \frac{\partial R}{\partial x} + m_3 \frac{\partial R}{\partial y} + n_3 \frac{\partial R}{\partial z}$$

gives

$$U = -\frac{k^2 M_d}{r_d^3} r \left[1 - 3\cos^2\beta + \frac{3}{2}\frac{r}{r_d}\left(3 - 5\cos^2\beta\right)\cos\beta \right]$$

$$V = \frac{3k^2 M_d}{r_d^3} r \left[\ell_2 \ell_d + m_2 m_d + n_2 n_d \right]\left[\cos\beta - \frac{1}{2}\frac{r}{r_d}\left(1 - 5\cos^2\beta\right)\right] \qquad (26\text{-}13)$$

$$W = \frac{3k^2 M_d}{r_d^3} r \left[\ell_3 \ell_d + m_3 m_d + n_3 n_d \right]\left[\cos\beta - \frac{1}{2}\frac{r}{r_d}\left(1 - 5\cos^2\beta\right)\right]$$

Then using equations (26-5) and (26-7) gives

$$\ell_2 \ell_d + m_2 m_d + n_2 n_d = -A \sin u + B \cos u \qquad (26\text{-}14)$$

and similarly (but by no means obviously), (26-6) and (26-7) give

$$\ell_3 \ell_d + m_3 m_d + n_3 n_d = \sin i \left[\cos u_d \sin(\Omega - \Omega_d) - \cos i_d \sin u_d \cos(\Omega - \Omega_d)\right]$$

$$+ \cos i \sin i_d \sin u_d = C \text{ (say)} \qquad (26\text{-}15)$$

Substituting equations (26-10), (26-14) and (26-15) into (26-13), the components of the disturbing force become

$$U = -\frac{k^2 M_d}{r_d^3} r \left[1 - \frac{3}{2}(A^2+B^2) - 3AB\sin 2u - \frac{3}{2}(A^2-B^2)\cos(2u) + \frac{3}{2}\frac{r}{r_d}(A\cos u + B\sin u)(3 - 5(A\cos u + B\sin u)^2) \right]$$

$$V = \frac{3k^2 M_d}{r_d^3} r \left[AB\cos 2u - \frac{1}{2}(A^2-B^2)\sin 2u + \frac{1}{2}\frac{r}{r_d}(A\sin u - B\cos u)\left\{1 - 5(A\cos u + B\sin u)^2\right\} \right] \quad (26\text{-}16)$$

$$W = \frac{3k^2 M_d}{r_d^3} r C \left[A\cos u + B\sin u - \frac{1}{2}\frac{r}{r_d}\left\{1 - 5(A\cos u + B\sin u)^2\right\} \right]$$

where the error due to the neglected terms is a factor $1 + 0\left(\frac{r^2}{r_d^2}\right)$.

Cook mentions that it is worth noting that A, B and C are direction cosines of the radius vector to the disturbing body referred to the geocentric axes through the ascending node of the satellite; through the apex of the orbit; and normal to the orbit respectively.

The equation (26-16) may now be substituted into Gauss's planetary equations and integrated. Before doing this we first change the independent variables to f, the true anomaly. Recall (22-17),

$$\frac{df}{dt} = \frac{\sqrt{\mu p}}{r^2} - \frac{d\omega}{dt} - \frac{d\Omega}{dt}\cos i \tag{26-17}$$

The main change in Ω and ω are those due to the oblateness factor.

Now

$$\frac{d\Omega_s}{dt}\cos i_0 = -\frac{3n_0 J_2 a_e^2 \cos^2 i_0}{2a_0^2(1-e_0^2)^2}, \quad \frac{d\omega_s}{dt} = \frac{3n_0 J_2 a_e^2}{a_0^2(1-e_0^2)^2}\left(\frac{5}{4}\sin^2 i - 1\right) \tag{26-18}$$

and since $\frac{d\Omega_s}{dt}$ is of the order of $10^{-3}\cos i_0$ and $\frac{d\omega_s}{dt}$ is the order 10^{-3} times $[\frac{5}{4}\sin^2 i - 1]$ for the oblateness effects, and since the luni-solar perturbations are of the order of 10^{-5}, we can approximate equation (26-17) as

$$\frac{df}{dt} = \frac{\sqrt{\mu p}}{r^2} = \frac{\sqrt{\mu}\,(1+e\cos f)^2}{\sqrt{a^3(1-e^2)^3}} \tag{26-19}$$

which is precisely our previous $df^{(o)}$ representation.

We can now use equation (26-19) to change the independent variable in (24-8) from t to f. Then substituting from (26-16) gives the following equations.

$$\frac{da}{df} = \frac{2Kr^3}{n^2a^2(1-e^2)}\left\{\left[-1+\frac{3}{2}(A^2+B^2)+3B\sin 2u+\frac{3}{2}(A^2+B^2)\cos 2u\right]e\sin f\right.$$
$$\left.+3\left[AB\cos 2u-\frac{1}{2}(A^2-B^2)\sin 2u\right]\left[1+e\cos f\right]\right\} \tag{26-20}$$

$$\frac{de}{df} = \frac{K r^3}{n^2 a^3}\left\{\left[-1+\frac{3}{2}(A^2+B^2)+3AB\sin 2u+\frac{3}{2}(A^2-B^2)\cos 2u\right]\sin f\right.$$

$$\left. + 3\left[AB\cos 2u - \frac{1}{2}(A^2-B^2)\sin 2u\right]\left[\cos f + \frac{\cos f + e}{1+e\cos f}\right]\right.$$

$$\frac{d\Omega}{df} = \frac{3Kr^4C}{n^2a^4(1-e^2)\sin i}\left[A\cos u + B\sin u\right]\sin u$$

$$\frac{di}{df} = \frac{3Kr^4C}{n^2a^4(1-e^2)}\left[A\cos u + B\sin u\right]\cos u$$

$$\frac{d\omega}{df} = \frac{K r^3}{n^2 a^3 e}\left\{\left[1-\frac{3}{2}(A^2+B^2)-3AB\sin u-\frac{3}{2}(A^2-B^2)\cos 2u\right]\cos f\right.$$

$$\left. + 3\left[AB\cos 2u - \frac{1}{2}(A^2-B^2)\sin 2u\right]\left[1+\frac{1}{1+e\cos f}\right]\sin f - \frac{d\Omega}{dt}\cos i \right. .$$

$$K = \frac{k^2 M_d}{r_d^3}$$

The rate of change of any orbital element can be integrated over one complete revolution to find its secular change. For example, integrating $\Delta\Omega$ over one revolution gives

$$\Delta\Omega = \frac{3K(1-e^2)^3 C}{2n^2 \sin i}\int_0^{2\pi}\frac{A\sin 2\omega\cos 2f + B(1-\cos 2\omega\cos 2f)}{(1+e\cos f)^4}\,df$$

$$\Delta\Omega = \frac{3\pi K C}{2n^2(1-e^2)^{1/2}\sin i}\left[5Ae^2\sin 2\omega + B(2+3e^2-5e^2\cos 2\omega)\right] .$$

This can be done for each of the elements to give the following secular changes in one period.

$$\frac{\Delta a}{P} = 0 + \text{Order}\left[\frac{k^2 M_d a^2 e}{n r_d^4}\right] = 0 + O(\Gamma a e) \tag{26-21}$$

$$\frac{\Delta e}{P} = -\frac{15 K e\sqrt{1-e^2}}{2n}\left[AB\cos 2\omega - \frac{1}{2}(A^2-B^2)\sin 2\omega\right] + O(\Gamma) \tag{26-22}$$

$$\frac{\Delta r_p}{P} = \frac{15 K a e\sqrt{1-e^2}}{2n}\left[AB\cos 2\omega - \frac{1}{2}(A^2-B^2)\sin 2\omega\right] + O(\Gamma a) \tag{26-23}$$

$$\frac{\Delta \Omega}{P} = \frac{3KC}{4n\sqrt{1-e^2}\,\sin i}\left[5Ae^2\sin 2\omega + B(2+3e^2-5e^2\cos 2\omega)\right] + O(\Gamma) \tag{26-24}$$

$$\frac{\Delta i}{P} = \frac{3KC}{4n\sqrt{1-e^2}}\left[A(2+3e^2+5e^2\cos 2\omega) + 5Be^2\sin 2\omega\right] + O(\Gamma) \tag{26-25}$$

$$\frac{\Delta \omega}{P} = \frac{3K\sqrt{1-e^2}}{2n}\left[5\left\{AB\sin 2\omega + \frac{1}{2}(A^2-B^2)\cos 2\omega\right\} - 1 + \frac{3}{2}(A^2+B^2)\right.$$
$$\left. + \frac{5a}{2er_p}\left\{1-\frac{5}{4}(A^2+B^2)\right\}\left\{A\cos\omega + B\sin\omega\right\}\right] + O\left[\frac{\Gamma a}{e r_d}, \frac{\Gamma a e}{r_p}\right] \tag{26-26}$$

$$\Gamma = \frac{Ka}{n r_d} = \frac{k^2 M_d a}{n r_d^4} \quad ; \quad K = \frac{k^2 M_d}{r_d^3} \ . \tag{26-27}$$

For circular orbits the only elements of interest are Ω and i which become

$$\frac{\Delta\Omega}{P} = \frac{3KBC}{2n\sin i} \quad ; \quad \frac{\Delta i}{P} = \frac{3KAC}{2n} \ . \tag{26-28}$$

There are also cases of possible resonances. When resonance occurs, the eccentricity is the most important element since any change in it affects the perigee radius, which influences the satellites' lifetime. After substituting for A and B from (26-11) and (26-22), Cook finds after considerable trigonometric manipulations that (26-22) reduces to the following

$$
\begin{aligned}
\frac{\Delta e}{p} = -\frac{15K}{4n} e\sqrt{1-e^2} \Bigg[& \sin^4\frac{i}{2} \Big\{ \cos^4\frac{i_d}{2} \sin 2(\Omega-\Omega_d-u_d-\omega) \\
& + \sin^4\frac{i_d}{2} \sin 2(\Omega-\Omega_d+u_d-\omega) + \frac{1}{2}\sin^2 i_d \sin 2(\Omega-\Omega_d-\omega) \Big\} \\
& - \cos^4\frac{i}{2} \Big\{ \cos^4\frac{i_d}{2} \sin 2(\Omega-\Omega_d-u_d-\omega) + \sin^4\frac{i_d}{2} \sin 2(\Omega-\Omega_d+u_d+\omega) \\
& + \frac{1}{2}\sin^2 i_d \sin 2(\Omega-\Omega_d+\omega) \Big\} + \cos^2\frac{i}{2} \sin i \sin i_d \Big\{ \cos i_d \sin(\Omega-\Omega_d+2\omega) \\
& + \cos^2\frac{i_d}{2} \sin(2u_d-\Omega-\Omega_d-2\omega) + \sin^2\frac{i_d}{2} \sin(\Omega-\Omega_d+2u_d+2\omega) \Big\} \\
& + \sin^2\frac{i}{2} \sin i \sin i_d \Big\{ \cos i_d \sin(\Omega-\Omega_d-2\omega) + \cos^2\frac{i_d}{2} \sin(2u_d-\Omega+\Omega_d+2\omega) \\
& + \sin^2\frac{i_d}{2} \sin(\Omega-\Omega_d+2u_d-2\omega) \Big\} - \frac{3}{8}\sin^2 i \sin^2 i_d \Big\{ \sin 2(\omega+u_d) + \sin 2(\omega-u_d) \Big\} \\
& - \frac{1}{2}\sin^2 i \Big\{ 1 - \frac{3}{2}\sin^2 i_d \Big\} \sin 2\omega \Bigg]
\end{aligned}
\tag{26-29}
$$

Clearly celestial mechanicians are men of persistent patience who think high thoughts and lead clean lives.

Fifteen possible cases when resonance can occur are:

$$
\begin{aligned}
&\dot{\Omega} - \dot{\Omega}_d \pm \dot{u}_d \pm \dot{\omega} = 0 \\
&2\dot{u}_d \pm \dot{\Omega} \mp \dot{\Omega}_d \pm 2\dot{\omega} = 0 \\
&\dot{\Omega} - \dot{\Omega}_d \pm \dot{\omega} = 0 \\
&\dot{\Omega} - \dot{\Omega}_d \pm 2\dot{\omega} = 0 \\
&\dot{\omega} \pm \dot{u}_d = 0 \\
&\dot{\omega} = 0
\end{aligned}
\tag{26-30}
$$

For any one of these conditions, equation (26-29) indicates we will have at least one constant term, i.e., a secular change in e. To determine if resonance occurs, that is, if any of the terms of (26-30) hold, one must consider all perturbing influences on terms like $\dot{\Omega}$, $\dot{\omega}$, etc. Near resonance conditions are also important.

For most purposes the only perturbing influence one need consider is the effect of the earth's gravitational oblateness where we can use the following:

$$\dot{\Omega} = -10.0 \left(\frac{a_e}{a}\right)^{3.5} \frac{\cos i}{(1-e^2)^2} \quad \text{in degrees per day}$$

$$\dot{\omega} = 5.0 \left(\frac{a_e}{a}\right)^{3.5} \frac{(5 \cos^2 i-1)}{(1-e^2)^2} \quad \text{in degrees per day.}$$

If we assume $\dot{\Omega}_d = 0$ for the moon then five of the relations in (26-30) produce resonance for both solar and lunar perturbations if the orbital inclination takes precisely one of the values given below.

Condition	Resonance inclination in degrees	
$\dot{\Omega} + \dot{\omega} = 0$	i = 46.4 or 106.8	
$\dot{\Omega} - \dot{\omega} = 0$	i = 73.2 or 133.6	
$\dot{\Omega} + 2\dot{\omega} = 0$	i = 56.1 or 111.0	(26-31)
$\dot{\Omega} - 2\dot{\omega} = 0$	i = 69 or 123.9	
$\dot{\omega} = 0$	i = 63.4 or 116.6	

For lunar perturbation, five additional cases of (26-30) can never occur and the remaining five are only possible in limited ranges of the inclination angle. The values of $\frac{a}{a_e}(1-e^2)^{4/7}$ which give resonance for particular values of the inclination are given in Figure 26-3. For solar perturbations, the values of $\frac{a}{a_e}(1-e^2)^{4/7}$ are given in Figure 26-4.

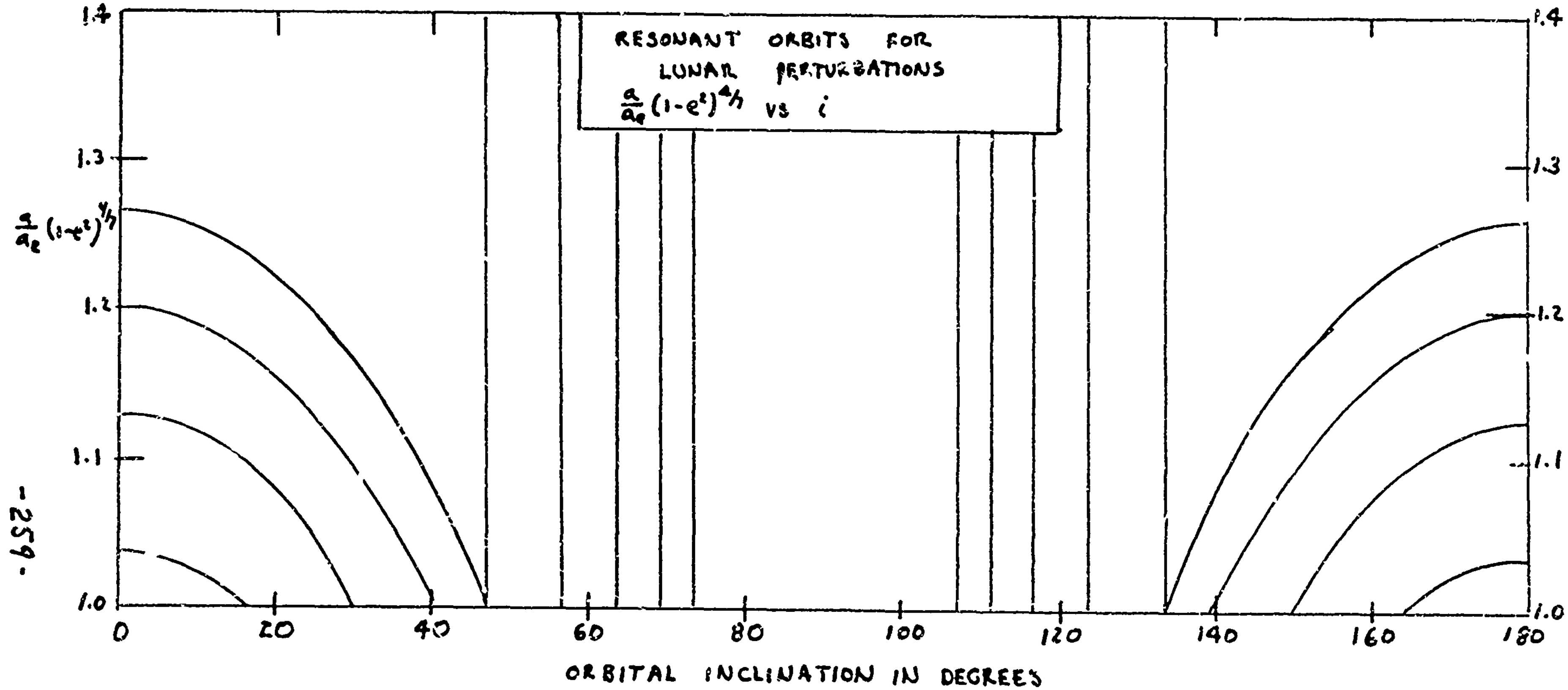

RESONANT ORBITS FOR
LUNAR PERTURBATIONS
$\frac{a}{a_e}(1-e^2)^{4/7}$ vs i
1.4
1.3
1.2
1.1
1.0
$\frac{a}{a_e}(1-e^2)^{4/7}$
0
20
40
60
80
100
120
140
160
180
ORBITAL INCLINATION IN DEGREES

FIGURE 26-3

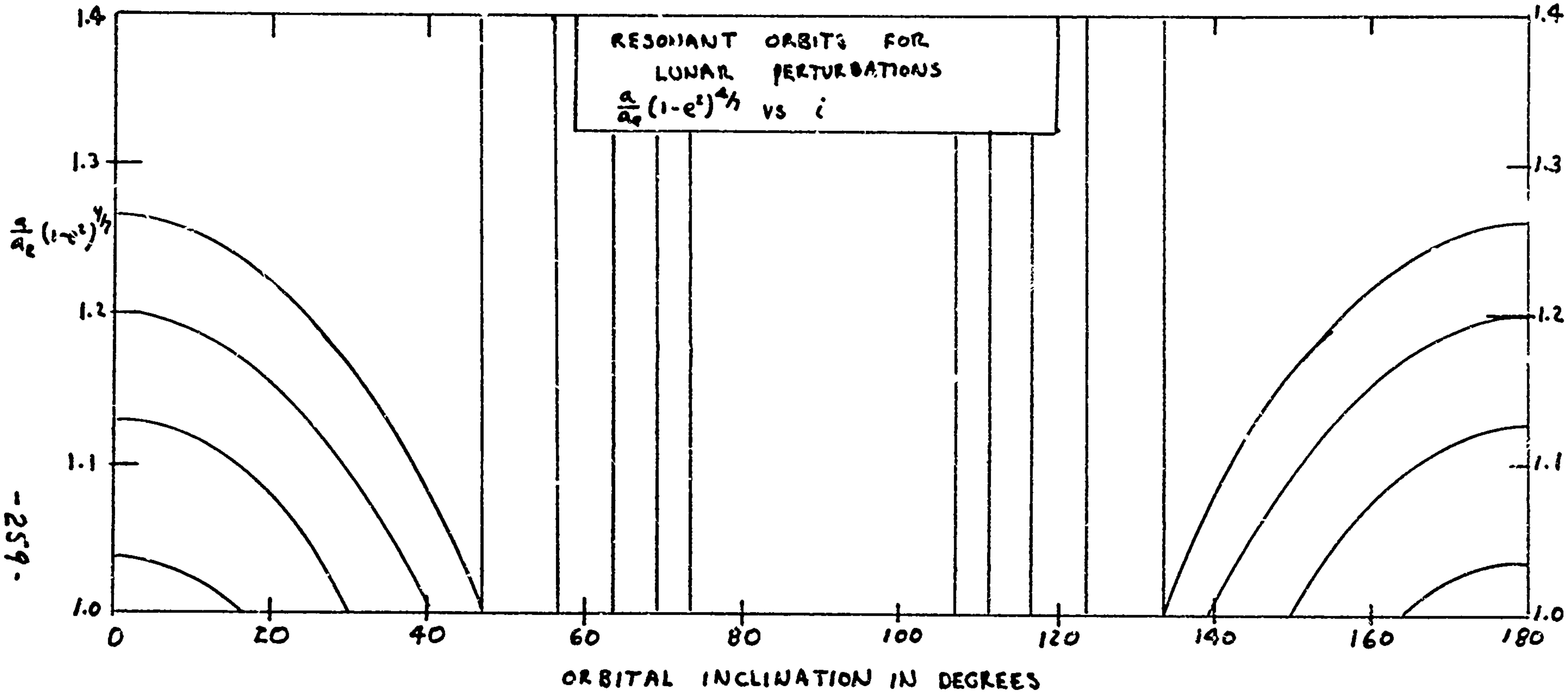

RESONANT ORBITS FOR LUNAR PERTURBATIONS
$\frac{a}{a_e}(1-e^2)^{4/7}$ vs i
1.4
1.3
1.2
1.1
1.0
$\frac{a}{a_e}(1-e^2)^{4/7}$
0
20
40
60
80
100
120
140
160
180
ORBITAL INCLINATION IN DEGREES

FIGURE 26-3

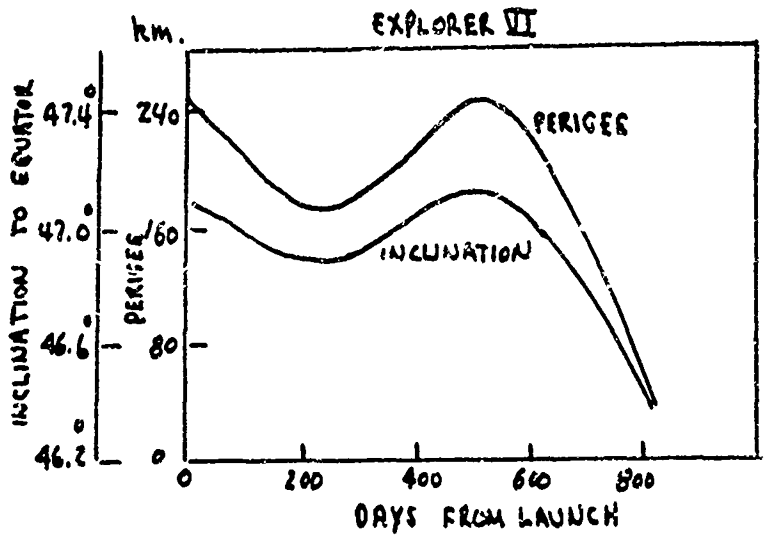


FIGURE 26-5 EXPLORER VI SATELLITE. (SMITH, 8).

The first case of resonance for earth satellites occurred with Explorer VI ($a_o = 4.34956a_e$, $e_o = 0.76121$, $i_o = 47.1°$, $\omega_o = 38.133°$, $\Omega_o = 59.205°$). Figure 26-5 shows how the satellite orbit changed with time. The original estimate of its lifetime was about 200 years, but Kozai showed that because of resonance it would be only the order of two years! By choosing a different time of day for the launch the solar and lunar perturbations on the orbit would be different and a longer lifetime can be achieved.

For very eccentric satellites, such as the Interplanetary Monitoring Platform (IMP), those perturbations cause large variations in the eccentricity and inclination angle. Figure 26-6 shows the results obtained by Barbara Shute. The inclination was also affected by changes as much as 20 degrees. The semi-major axis and mean motion of the satellite are not effected.

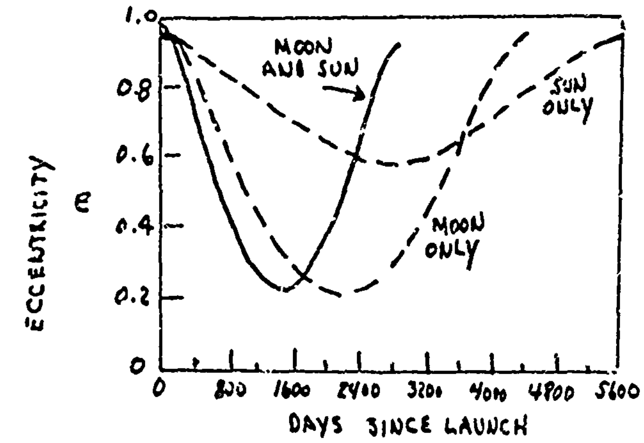


FIGURE 26-6 ECCENTRICITY VS TIME FOR IMP SATELLITE DUE TO SOLAR AND LUNAR PERTURBATIONS (SHUTE REF. 7).

Because of the eccentricity changes there is a substantial effect on perigee height of the satellite. Figure 26-7 shows the behavior of the perigee height as a function of time for an orbit with elements a = 50,000 km., e = 0.867, i = 33.0°, ω = 153.5° with the parameter of the curves being Ω. This latter can be considered as the launch time, all other things being constant. Thus one hour change in launch time corresponds to each 15° of Ω. This phenomena of perigee pumping can be used to raise the perigee when using a low energy booster or to cause the satellite to re-enter the atmosphere after a predetermined time.

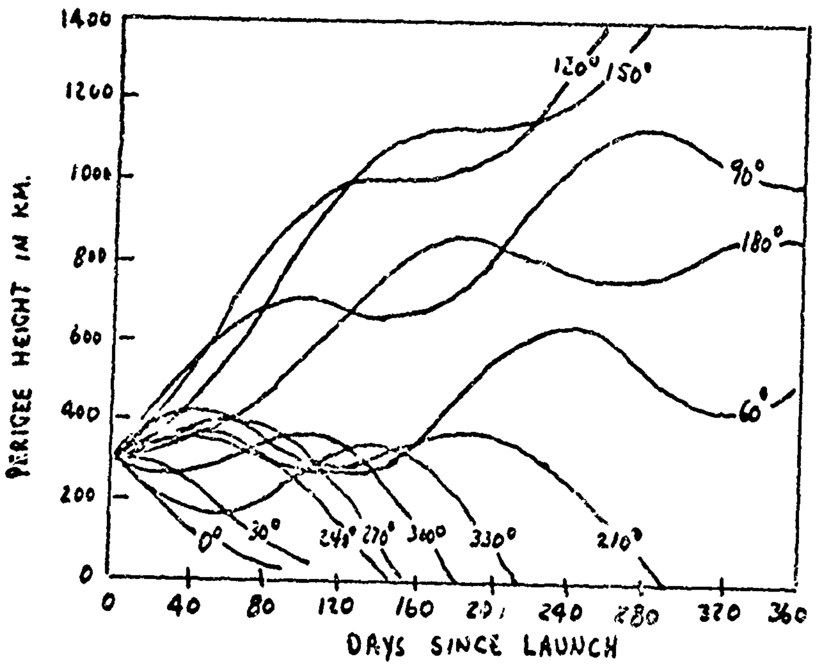


FIGURE 26-7 LUNI-SOLAR PERTURBATION EFFECT ON PERIGEE HEIGHT VS TIME. PARAMETER IS RIGHT ASCENSION OF ASCENDING NODE.

How well one can accomplish this depends on how closely these injection parameters and launch time can be controlled.

Because of these resonance problems we have to resort to numerical integration for very eccentric orbits. This is usually carried out together with some method, usually numerical, of removing the short period terms. One such method is due to Halphen which is discussed in detail in the references of Musen (5,6), of Smith (8) and of Shute (7).

REFERENCES

1. Kozai, Y., "On the effects of the Sun and the Moon upon the motion of a close Earth satellite," Smithsonian Astrophys. Observatory Special Report No. 22, pages 7-10, March 1959.

2. Cook, G. E., "Luni-Solar perturbations of the orbit of an Earth satellite," RAE (Farnborough) Tech Note GW-582, July 1961 (AD-266644).

3. Geyling, F. T., "Satellite perturbations from extraterrestrial gravitation and radiation pressure," Journal of Franklin Institute, Vol 269, page 375, 1960.

4. Moe, M. M., "Solar-lunar perturbations of the orbit of an earth satellite," ARS Journal, Vol 30, page 485, May 1960.

5. Musen, P., "Contributions to the theory of satellite orbits" Space Research Vol I, North Holland Publishing Comp., Amsterdam pages 434-447, 1960.

6. Musen, P., A. Bailie, E. Upton, "Development of the lunar and solar perturbations in the motion of an artificial satellite" NASA TN D-494, January 1961.

7. Shute, B., "Prelaunch analysis of high eccentricity orbits" NASA Technical Note D-2530, 1964.

8. Smith, A. J., "A discussion of Halphen's method for secular perturbations and its application to the determination of long range effects in the motion of celestial bodies" Part II, NASA Technical Report R-194, 1964.

27. SOLAR RADIATION PERTURBATIONS

Light quanta emanating from the sun carry a momentum equal to $\frac{h\nu}{c}$ where h is Planck's constant and ν is the frequency of the light. When the light impinges on a surface all or part of this momentum is imparted to the surface depending on the spectral characteristics of the surface. The existence of radiation pressure was first proclaimed by James C. Maxwell in 1871. Experimental evidence of its existence was first demonstrated by Lebedew in 1900. The force exerted by radiation impinging upon a satellite depends upon the intensity of the radiation, the presented normal area, and the reflectivity of the surface. The energy output of the sun is not constant in time, and there are periods of solar disturbances when high activity is reported. For purposes of this section, a constant mean value will be assumed.

For a non-spherical satellite, the magnitude of the force will depend on the satellite's orientation, but again we assume that some suitable average value, F per unit mass, acts while the satellite is in sunlight. Further, since the earth's distance from the sun is large compared with the size of the orbit, the force produced on a satellite by solar radiation pressure can be assumed independent of its <u>distance</u> from the sun. With these assumptions, the magnitude of the perturbing acceleration on a particle of mass m and cross sectional area A becomes

$$F = \frac{(1+\rho)S}{C} \left(\frac{A}{m}\right) \text{ cm/sec}^2 \tag{27-1}$$

where

ρ = the reflectivity, varying between 0 and 1

S = solar constant ($\approx 1.35 \times 10^6$ ergs/cm^2 sec)

C = velocity of light ($\approx 3 \times 10^{10}$ cm/sec)

It is sometimes convenient to express the force in terms of g's (1 g = 980 cm/sec^2). We then have

$$\frac{F}{g} = 4.5\ (1+\rho)\ \frac{A}{m} \tag{27-2}$$

For Vanguard I:

m = 1459 grams; A = 207.6 cm^2; $\frac{A}{m}$ = 0.1425 cm^2/g; F $\simeq 1.28 \times 10^{-5}$ cm/sec^2.

For Echo I:

m = 71215 grams; A = 7,292,901 cm^2; $\frac{A}{m}$ = 102 cm^2/g; F $\simeq 0.92 \times 10^{-2}$ cm/sec^2.

Using the notation of section 26, this radiation force F can be resolved into three components as

$$U = -F\ (A \cos u + B \sin u) \tag{27-3}$$

$$V = -F\ (-A \sin u + B \cos u) \tag{27-4}$$

$$W = -FC \tag{27-5}$$

These may be substituted into the Gauss planetary equations (24-8). One then transforms the independent variable from t to f as in section 26.

As an example consider the semi-major axis. After substituting (27-3) and (27-4) in (24-8) and changing to f as the independent variable, one has

$$\frac{da}{df} = \frac{2F(1-e^2)}{n^2}\left[\frac{A\sin u - B\cos u + e\,(A\sin\omega - B\cos\omega)}{(1 + e\cos f)^2}\right]. \qquad (27\text{-}6)$$

If the true anomaly when the satellite departs from and enters the earth's shadow are denoted by f_d and f_e respectively, the change in 'a' during one revolution is given by

$$\Delta a = \frac{2F(1-e^2)}{n^2}\int_{f_d}^{f_e} \frac{(A\sin\omega - B\cos\omega)(e+\cos f)+(A\cos\omega+B\sin\omega)\sin f}{(1+e\cos f)^2}\,df. \qquad (27\text{-}7)$$

Evaluating the integral and using the equation of the orbit gives,

$$\Delta a = \frac{2}{n^2 a}\left[(r_e\sin f_e - r_d\sin f_d)\,T_p + a\,(\cos E_e - \cos E_d)S_p\right] \qquad (27\text{-}8)$$

where

$$\begin{aligned} T_p = F\Big\{&\Big[\cos^2\frac{\varepsilon}{2}\sin(\omega+\Omega-L) + \sin^2\frac{\varepsilon}{2}\sin(\omega+\Omega+L)\Big]\cos^2\frac{i}{2} \\ &+ \Big[\cos^2\frac{\varepsilon}{2}\sin(\omega-\Omega-L) + \sin^2\frac{\varepsilon}{2}\sin(\omega-\Omega-L)\Big]\sin^2\frac{i}{2} \qquad (27\text{-}9) \\ &- \frac{1}{2}[\sin(\omega+L) - \sin(\omega-L)]\sin i\sin\varepsilon\Big\}, \end{aligned}$$

$$\begin{aligned} S_p = -F\Big\{&\Big[\cos^2\frac{\varepsilon}{2}\cos(\omega+\Omega-L) + \sin^2\frac{\varepsilon}{2}\cos(\omega+\Omega+L)\Big]\cos^2\frac{i}{2} \\ &+ \Big[\cos^2\frac{\varepsilon}{2}\cos(\omega-\Omega+L) + \sin^2\frac{\varepsilon}{2}\cos(\omega-\Omega-L)\Big]\sin^2\frac{i}{2} \qquad (27\text{-}10) \\ &+ \frac{1}{2}[\cos(\omega-L) - \cos(\omega+L)]\sin i\sin\varepsilon\Big\}, \end{aligned}$$

where ε is the obliquity (see Chapter 32) and $L = f + \omega$ for the sun. In terms of Chapter 26 we have evaluated A and B of equations (26-11) in equations (27-3) to (27-5) by using $i_d = \varepsilon$, $\Omega_d = 0$ and $u_d = L$ for the sun.

The other orbital elements are obtained by Cook. The problem now remains to determine the time at which f_e and f_d occur. This is further complicated by the oblateness effects which cause precession of the orbital elements Ω and ω. Thus a complete satisfactory analytical solution including passage through the earth's shadow has not been formulated. The best method appears to be successive applications of Cook's formulae on a piecewise basis or to do what has been done in the past, evaluate the Gauss planetary equations by numerical integration.

Escobal gives a method for determining the orbital entrance and exist of a satellite from the shadow of the earth [Escobal, pages 155-162].

If one wishes to be more accurate, the transient, slow, and diurnal fluctuations in F must be introduced; the radiation pressure reflected from the earth must be included and the disturbing force of the earth's shape and atmosphere must be properly contained. In spite of the sad state of any exact treatment, the analysis to date has been remarkably effective in accounting for the observed disturbances in artificial satellite orbits. Figure 27-1 shows some results given by R. K. Squires which compare the actual observed perigee height for Echo I with that calculated by numerical integration of the complete equations including radiation pressure and aerodynamic drag. Considering that numerical integration

programs are subject to accumulated effects of round-off error, the agreement is very good. The uncertainty in radiation reflectivity and drag are too large to warrant much further refinement of the calculations.

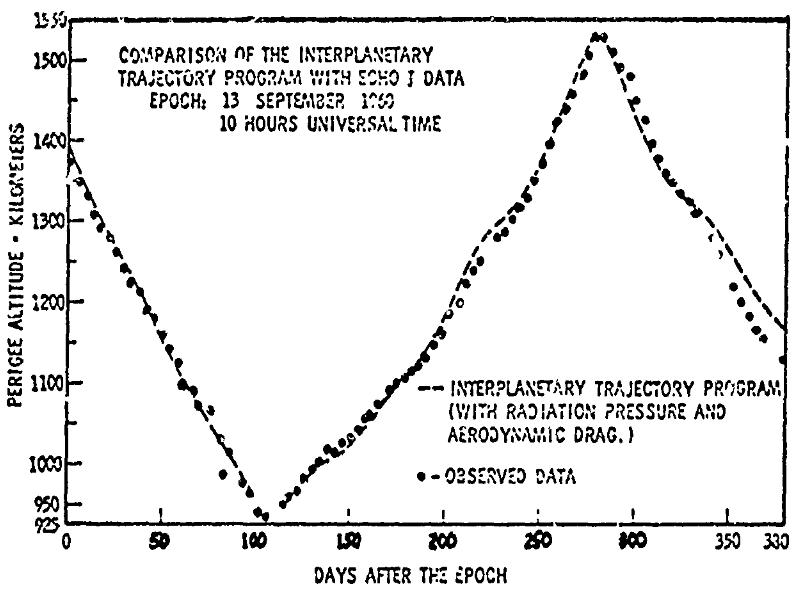


Figure 27-1 Perigee altitude vs time in orbit. (From R. K. Squires).

If the satellite is permanently in the sunlight the integrals like (27-7) are taken from 0 to 2π and the results simplify considerably.

$$\frac{\Delta a}{P} = 0$$

$$\frac{\Delta e}{P} = \frac{3\sqrt{1-e^2}}{2na}\, T_p$$

$$\frac{\Delta r_p}{P} = -\, a\, \frac{\Delta e}{P} = -\, \frac{3\sqrt{1-e^2}}{2n}\, T_p \qquad (27\text{-}11)$$

$$\frac{\Delta \Omega}{P} = -\, \frac{3We \sin \omega}{2na \sqrt{1-e^2} \sin i}$$

$$\frac{\Delta i}{P} = -\, \frac{3We \cos \omega}{2na \sqrt{1-e^2}}$$

$$\frac{\Delta \omega}{P} = -\, \frac{3\sqrt{1-e^2}}{2nae}\, S_p + \frac{3We \sin \omega \cot i}{2na\sqrt{1-e^2}}$$

with

$$W \sin \omega = \frac{F}{2} \Big\{ [\cos(\omega+\Omega-L) - \cos(\omega-\Omega+L)] \sin i \cos^2 \frac{\varepsilon}{2}$$

$$+ [\cos(\omega+\Omega+L) - \cos(\omega-\Omega-L)] \sin i \sin^2 \frac{\varepsilon}{2} \qquad (27\text{-}12)$$

$$+ [\cos(\omega+L) - \cos(\omega-L)] \cos i \sin \varepsilon \Big\},$$

$$W \cos \omega = \frac{F}{2} \{ [\sin(\omega+\Omega-L) - \sin(\omega-\Omega+L)] \sin i \cos^2 \frac{\varepsilon}{2}$$

$$+ [\sin(\omega+\Omega+L) - \sin(\omega-\Omega-L)] \sin i \sin^2 \frac{\varepsilon}{2} \qquad (27\text{-}13)$$

$$+ [\sin(\omega+L) - \sin(\omega-L)] \cos i \sin \varepsilon \} .$$

When the orbit is entirely in the sunlight, there are six conditions for resonance. They are

$$\dot{\omega} \pm \dot{\Omega} \pm \dot{L} = 0$$

$$\dot{\omega} \pm \dot{L} = 0.$$

These conditions are the same as six of equations (26-30).

Considerable insight can be obtained by studying the case where the shadow is neglected. For a typical near earth satellite ($a = 8000$ km, $e = 0.1$) studied by R. R. Allan he found that the orbit elements that do not vanish when the earth's shadow is neglected are changed by less than 25%, while those quantities that do vanish when the shadow is neglected are small anyway. The effect is less for more distant orbits, even of high eccentricity. For Echo type satellites, if the shadow effect is neglected, there is a loss of several degrees of accuracy in the true anomaly after say 100 days of flight.

The essential difference between sunlit and shadow cases is that there is no long period or secular change of the semi-major axis except when the earth's shadow is included.

Eugene Levin in an Aerospace Corporation report develops some very simple planar models which beautifully illustrate the essential features of the solar radiation effect. He assumes the earth, satellite and sun all lie in the same plane. When he assumes a fixed coplanar sun, he shows that the perigee decreases linearly with time. When the sun was assumed to revolve in the orbit plane at a constant rate, the variation of perigee was changed from a linear decay to a periodic fluctuation and perigee position precessed at half the rate of the source. Oblateness of the earth has an effect on the perigee precession rate and this effect is strongly dependent on the orbital inclination. As an approximation, the perigee will oscillate with a period of $2\pi/|\dot{L}-\dot{\Omega}-\dot{\omega}|$ where $\dot{L}$ is the angular rate of the earth-sun line and $\dot{\Omega}$ and $\dot{\omega}$ are the oblateness precession rates for the nodal and argument of perigee angles. These latter two rates, of course, depend on the orbital inclination angle.

Using the same coplanar moving sun model, Levin also shows that the shadow affects the amplitude of the perturbation by an amount that is approximately equal to the fractional part of the orbit in shadow. He finds the reflected radiation from the earth to the satellite is negligible.

These simple models of Levin give a very keen insight into the mechanics of the radiation perturbing effect and are in surprisingly good agreement with more sophisticated methods.

In general, solar radiation pressure causes first order perturbations in all six orbital elements. For a circular orbit it displaces its geometric center perpendicular to the earth-sun line, in the orbit plane and in a direction so as to decrease the altitude of that part of the orbit in which the satellite moves away from the sun. The effect is dependent upon the orientation of the orbit and may even change sign as the orbit rotates. Again, resonant conditions must be carefully considered but the overall general effect is a "sideways" motion of the orbit.

References

1. Cook, G. E., "Luni-Solar Perturbations of the Orbit of an Earth Satellite," RAE (Farnborough) Tech Note GW 582, July 1961 (AD 266644).

2. Levin, E., "Orbit Perturbations Due to Solar Radiation Pressure," SSD-TR-65-97 (TDR-469 (5540-10) -7), Aerospace Corp., 28 June 1965 (AD-468419).

3. Polyakhava, Ye. N., "Solar Radiation Pressure and the Motion of Earth Satellites," AIAA Journal, Vol 1, No. 12, December 1963, pages 2893-2909. Translation of Bull. of Inst. of Ther. Astron. USSR.

4. Allan, R. R., "Satellite Orbit Perturbations Due to Radiation Pressure and Luni-Solar Forces," RAE (Farnborough) Technical Note No. Space 7, February 1962 (AD-275628).

5. Musen, P., "The Influence of Solar Radiation Pressure on the Motion of an Artificial Satellite", Journal of Geophysical Research, Vol. 65, page 1391, 1960.

6. Kozai, Y., "Effects of Solar Radiation Pressure on the Motion of an Artificial Satellite," Smithsonian Observatory, Special Report No. 56, pp 25-33, January 1961.

28. DRAG PERTURBATIONS

Because of the relative velocity between the satellite and the surrounding atmosphere, there is a resultant aerodynamic force acting on the satellite. This force is assumed to act through the center of gravity of the satellite in an arbitrary direction. The resultant force is decomposed along a set of mutually perpendicular axes. The three force components are:

1. Drag: This force component is in the direction of the relative velocity between satellite and atmosphere. Positive drag is opposite to the velocity direction.
2. Lift: When the drag is vectorially subtracted from the resultant force, the force component which remains and is perpendicular to the relative velocity is called the lift. It is the component in the plane of symmetry of the satellite.
3. Side Force: The force component that finally remains is the side force. This component is zero in symmetric flight where the velocity vector lies in the plane of symmetry of the satellite.

The lift and drag forces are proportional to the dynamic pressure $q = \frac{1}{2}\rho\tilde{v}^2$ and a reference area A, and the constants of proportionality which are respectively the lift and drag coefficients C_L and C_D.

$$\frac{D}{m} = \frac{1}{2}\frac{C_D A}{m}\rho\tilde{v}^2 \tag{28-1}$$

$$\frac{L}{m} = \frac{1}{2}\frac{C_L A}{m}\rho\tilde{v}^2 = \frac{1}{2}\frac{C_L}{C_D}\frac{C_D A}{m}\rho\tilde{v}^2 = \frac{1}{2}\frac{L}{D}\frac{C_D A}{m}\rho\tilde{v}^2 \quad . \tag{28-2}$$

ρ is the atmospheric density and $\tilde{v}$ is the satellite velocity relative to the ambient air. Let

$$B = \frac{C_D A}{m} \tag{28-3}$$

$$\frac{C_L A}{m} = \frac{C_L}{C_D} B = \frac{L}{D} B$$

Where $\frac{L}{D}$ is the lift to drag ratio and B is called the ballistic coefficient. For our calculations we need the force in terms of v, the velocity of the satellite relative to the center of the earth. If we use

$$\delta = \frac{\tilde{v}^2}{v^2} \tag{28-4}$$

then the drag and lift forces may be written

$$\frac{D}{m} = \frac{1}{2\rho} v^2 \delta B \; ; \qquad \frac{L}{m} = \frac{1}{2} \rho v^2 \frac{L}{D} \delta B \tag{28-5}$$

T. E. Sterne, in his article in the Journal of the American Rocket Society, Vol. 29 (1959), page 777, shows that δ may be represented by

$$\delta = \left(1 - \frac{r_p \omega_e}{v_p} \cos i\right)^2 \tag{28-6}$$

(See also Deutsch's book, pages 208-209)

$r_p = a(1-e)$ is the perigee measured from the earth's center. The term ω_e should really be the angular velocity of the atmosphere at perigee but observations indicate this is practically equal to the earth's angular velocity ω_e. For heights of 200–250 km the correct value does not differ from ω_e by more than 1.2 at the very most. The value of δ for an eastbound satellite ($i < 90°$) usually lies between 0.9 and 1.0 so we shall consider it a constant nearly equal to unity.

In order to calculate the density, $\rho(r)$, we shall study the effect of drag on the orbital parameters. The gravity anomaly effects on the Vanguard

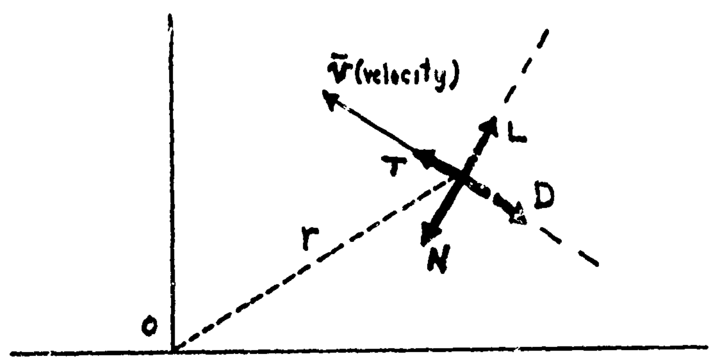


Figure 29-1

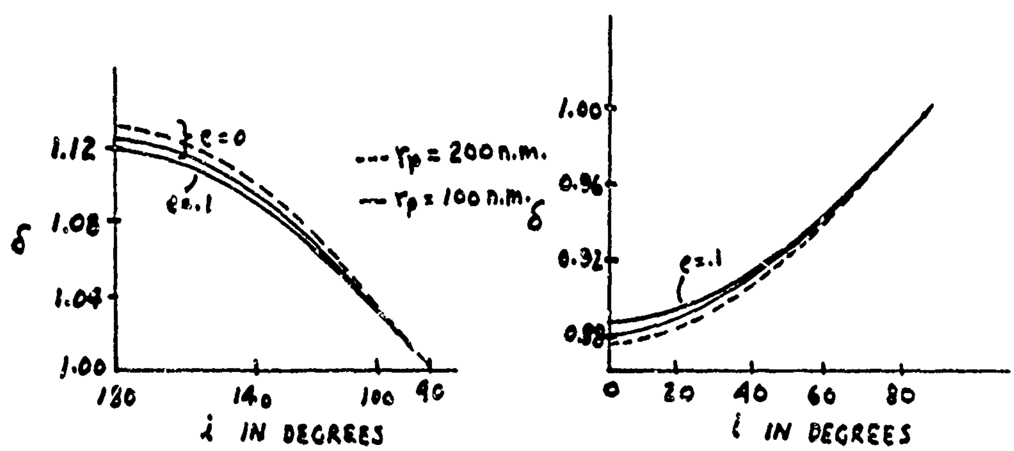


Figure 29-2

satellite are of the order of one dyne while the drag forces are of the order of 10^{-5} dynes. Both forces are small indeed but they have measurable accumulative effects. The oblateness terms do not effect the semi-major axis and hence do not give rise to a secular change in the period ($n^2a^3 = \mu$, $P = \frac{2\pi}{n}$). Thus by measuring ΔP we can hopefully calculate $\rho(r)$.

First let us digress and consider the composition of the drag force. The drag and lift forces are in the -T and -N force directions (see figure 24-2) and hence

$$T = -\frac{D}{m} = -\frac{1}{2} B \delta \rho(r) \mathbf{V}^2$$

$$N = -\frac{L}{m} = -\frac{1}{2}\frac{L}{D} B \delta \rho(r) \mathbf{V}^2 \tag{28-7}$$

$$W = 0 .$$

Let's now develop an expression for the velocity $\mathbf{V}$ as a function of the orbital element.

The total energy is given by

$$E = -\frac{\mu}{2a} = \frac{1}{2}\mathbf{V}^2 - \frac{\mu}{r} \tag{28-8}$$

hence

$$\mathbf{V}^2 = \mu\left(\frac{2}{r} - \frac{1}{a}\right) \tag{28-9}$$

which is the vis-viva integral (equation 2-30).

$$\mathbf{V}^2 = \mu\left[\frac{2a - r}{ar}\right] \tag{28-10}$$

letting $r = \frac{a(1-e^2)}{1 + e\cos f}$ and reducing gives

$$v^2 = \frac{\mu}{a(1-e^2)}\,[1 + 2e\cos f + e^2] \tag{28-11}$$

or if we let $r = a\,(1-\cos E)$ and substitute into equation (28-10) we obtain

$$v^2 = \left(\frac{\mu}{a}\right)\frac{1 + e\cos E}{1 - e\cos E}\,. \tag{28-12}$$

These two formulae give the velocity of a satellite at any point around the orbit. From (28-11) we have the relation

$$1 + 2e\cos f + e^2 = \frac{v^2 a(1-e^2)}{\mu} \tag{28-13}$$

and using

$$p = a(1-e^2)\quad , \text{ we have} \tag{28-14}$$

$$1 + 2e\cos f + e^2 = \frac{v^2 p}{\mu}\,. \tag{28-15}$$

We can use this expression to simplify the set of equations (24-13).

Now the drag and lift forces (28-7) can be written using (28-12) as

$$T = -\frac{1}{2}\,\frac{B}{a}\,\delta\,\rho(r)\,\mu\,\frac{1 + e\cos E}{1 - e\cos E} \tag{28-16}$$

$$N = -\frac{1}{2}\,\frac{L}{aD}\,B\,\delta\,\rho(r)\,\mu\,\frac{1 + e\cos E}{1 - e\cos E}. \tag{28-17}$$

and for symmetrical satellite

$W = 0$

Now that we have the drag and lift forces let's return to the calculation of the change in the semi-major axis.

Given $\left(\frac{2\pi}{P}\right)^2 a^3 = \mu$ we can differentiate to obtain

$$3(2\pi)^2 a^2 \frac{da}{dt} = 2\mu P \frac{dP}{dt} \tag{28-18}$$

or

$$\frac{dP}{dt} = \frac{3P}{2a}\frac{da}{dt} \, . \tag{28-19}$$

From equations (24-13) we have

$$\frac{da}{dt} = \frac{2T\sqrt{1 + 2e \cos f + e^2}}{n\sqrt{1 - e^2}}$$

hence

$$\frac{dP}{dt} = \frac{3P}{2a}\frac{2T\sqrt{1 + 2e \cos f + e^2}}{n\sqrt{1 - e^2}} \tag{28-20}$$

The change in P over one revolution is given by

$$\frac{\Delta P}{P} = \frac{1}{P}\left[P(t_o + P) - P(t_o)\right] = 3\int_{t_o}^{t_o+P} \frac{T\sqrt{1 + 2e \cos f + e^2}}{an\sqrt{1 - e^2}}\, dt \tag{28-21}$$

If we substitute (28-16) for the T force in the integral of equation (28-21) it becomes

$$\frac{\Delta P}{P} = -\frac{3 B \delta \mu}{2n a^2\sqrt{1 - e^2}} \int_{t_o}^{t_o+P} \rho(r)\sqrt{1 + 2e \cos f + e^2} \left[\frac{1 + e \cos E}{1 - e \cos E}\right] dt \tag{28-22}$$

From equations (28-11) and (28-12) we have

$$1 + 2e \cos f + e^2 = \frac{V^2 a(1-e^2)}{\mu} = \frac{\mu}{a}\left[\frac{1 + e \cos E}{1 - e \cos E}\right]\frac{a(1-e^2)}{\mu}$$

so that

$$1 + 2e \cos f + e^2 = (1 - e^2)\,\frac{1 + e \cos E}{1 - e \cos E}\ . \tag{28-23}$$

Using this, (28-22) may be written as

$$\frac{\Delta P}{P} = -\frac{3B\ \delta\ \mu}{2na^2}\int_{t_o}^{t_o+P} \rho(r)\,\frac{(1 + e \cos E)^{3/2}}{(1 - e \cos E)^{3/2}}\,dt\ . \tag{28-24}$$

To integrate we change the variable of integration to dE by using the same first order approximation as when we changed from dt to df. Since

$$nt - n\tau_o = E - e \sin E \tag{28-25}$$

we can write

$$n_o dt = (1 - e_o \cos E^{(0)})dE^{(0)} \tag{28-26}$$

for the osculating ellipse and hence (28-24) can be written ($n = \sqrt{\mu}\, a^{-3/2}$)

$$\frac{\Delta P}{P} = -\frac{3B\delta}{2} a_o \int_0^{2\pi} \rho(r) \frac{(1 + e_o \cos E^{(0)})^{3/2}}{(1 - e_o \cos E^{(0)})^{1/2}} \, dE^{(0)} \qquad (28\text{-}27)$$

In order to proceed further we must make some assumptions about the density $\rho(r)$. If we assume a hydrostatic relation for the vertical distribution of air pressure then

$$dp = -g\rho dr \qquad (28\text{-}28)$$

p = pressure, g = acceleration of gravity, ρ = density

or

$$\frac{dp}{p} = -\frac{mg}{kT} dr = -\frac{1}{H} dr \qquad (28\text{-}29)$$

Where m = average particle mass, T = absolute temperature at height r, k = Boltzmann's constant and H = scale height.

If we assume isothermal, constant g conditions, H will be a constant and we can integrate (28-29) to give

$$p = p_{r_p} \exp\left\{-\frac{1}{H}(r - r_p)\right\} \qquad (28\text{-}30)$$

Where p_{r_p} is the pressure at some reference height $r = r_p$. The density relation then follows.

$$\rho = \rho_{r_p} \exp\left\{-\frac{1}{H}(r - r_p)\right\} \qquad (28\text{-}31)$$

If we had considered variation of gravity with altitude, $g = \frac{g_0 r_p^2}{(r_p + r)^2}$, then

$$\rho = \rho_{r_p} \exp\left\{ - \frac{m g_0 r_p r}{kT(r_p + r)} \right\}$$

but we shall neglect this for our discussions.

Now let $r = a(1-e \cos E)$ and $r_p = a(1-e)$ and we obtain

$$\rho = \rho_{r_p} \exp\left[-\frac{1}{H}\left(a - ae \cos E - a(1-e) \right) \right] \tag{28-32}$$

$$\rho = \rho_{r_p} \exp\left[- \frac{ae}{H}(1 - \cos E) \right] = \rho_{r_p} e^{-\frac{ae}{H}} e^{\frac{ae}{H}\cos E} \tag{28-33}$$

Now lets return to the osculating elements and substitute for $\rho(r)$ in the $\frac{\Delta P}{P}$ integral of equation (28-27).

$$\frac{\Delta P}{P} = -\frac{3}{2} a_0 B\delta \int_0^{2\pi} \rho_{r_p} e^{-\frac{a_0 e_0}{H}} e^{\frac{a_0 e_0}{H}\cos E^{(0)}} \left[\frac{(1 + e_0 \cos E^{(0)})^{3/2}}{(1 - e_0 \cos E^{(0)})^{1/2}} \right] dE^{(0)} \tag{28-34}$$

To calculate we first expand the bracket term in a series.

$$\frac{(1+e\cos E)^{3/2}}{(1-e\cos E)^{1/2}} \frac{\sqrt{1+e\cos E}}{\sqrt{1+e\cos E}} = \frac{(1+e\cos E)^2}{\sqrt{1-e^2\cos E}} = 1 + 2e\cos E + \frac{3}{2} e^2 \cos^2 E + e^3 \cos^3 E + \ldots .$$

hence -

$$\frac{\Delta \rho}{\rho} = -\frac{3}{2} a_o B \delta \int_0^{2\pi} \rho_{r_p} e^{-\frac{a_o e_o}{H}} e^{\frac{a_o e_o}{H}\cos E^{(o)}} \left\{ 1 + 2e_o \cos E^{(o)} + \frac{3}{2} e_o^2 \cos^2 E^{(o)} + e_o^3 \cos^3 E^{(o)} + \ldots \right\} dE^{(o)} .$$

Now

$$\int_0^{2\pi} e^{\frac{ae}{H}\cos E} dE = 2\pi I_o\left(\frac{ae}{H}\right) \qquad (28\text{-}35)$$

where $I_o\left(\frac{ae}{H}\right)$ is the Bessel function of order zero and purely imaginary argument.

$$I_o(c) = J_o(ic) = \frac{1}{\pi}\int_0^{\pi} e^{c\cos E} dE . \qquad (28\text{-}36)$$

By successively differentiating both sides with respect to c and using known relations between $I_o(c)$ and $I_1(c)$, one finds

$$\begin{aligned} \frac{1}{\pi}\int_0^{\pi} \exp\left[c\cos E\right] \cos^n E \, dE &= I_o(c) \quad , \text{ for } n=0 . \\ &= I_1(c) \quad , \quad n=1 . \\ &= I_o(c) - \frac{1}{c} I_1(c) \quad , \quad n=2 . \\ &= \frac{1}{c} I_o(c) + \left(1+\frac{2}{c^2}\right) I_1(c) \quad , \quad n=3 . \end{aligned} \qquad (28\text{-}37)$$

We can thusly evaluate each term of equation (28-34) to give the following:

$$\frac{\Delta \rho}{\rho} = 3\pi \, a_o \, B\delta \rho_{r_p} e^{-\frac{a_o e_o}{H}} \left\{ I_o\left(\frac{a_o e_o}{H}\right)\left[1 + \frac{3}{2}e_o^2 - \frac{H e_o^2}{a_o}\right] \right. \qquad (28\text{-}38)$$

$$\left. + I_1\left(\frac{a_o e_o}{H}\right)\left[2e_o - \frac{3He_o}{2a_o} + e_o^3 + \frac{2H^2 e_o}{a_o^2}\right] \right\}$$

which is correct through terms in e_0^3. When $\frac{a_o e_o}{H}$ is small, say $\frac{a_o e_o}{H} < 2$, we can expand the terms in the brackets { } to give

$$\left\{\cdots\right\} = 1 + \frac{3}{4}e_o^2 + \frac{a_o e_o^2}{H}\left(1 + \frac{3}{8}e_o^2\right) + \frac{1}{4}\frac{a_o^2 e_o^2}{H^2}\left(1 + \frac{9}{8}e_o^2\right)$$
$$+ \frac{a_o^3 e_o^4}{8H^3}\left(1 + \frac{5}{12}e_o^2\right) + \frac{a_o^4 e_o^4}{64H^4}\left(1 + \frac{5}{4}e_o^2\right) + \cdots \tag{28-39}$$

which reduces to unity as $e_o \to 0$.

When $\frac{a_o e_o}{H} > 2$, the asymptotic expansion for the Bessel functions may be used to give

$$\frac{\Delta P}{P} = -\frac{3}{2} a_o B\delta \rho_{r_p}\sqrt{\frac{\pi H}{2a_o e_o}}\left\{1 + \frac{H}{8a_o e_o}\left(\frac{1 - 8e_o + 3e_o^2}{1 - e_o^2}\right)\right.$$
$$\left. + \frac{9H^2}{128 a_o^2 e_o^2}\left(\frac{3 - 16e_o + 50e_o^2 + 16e_o^3 - 5e_o^4}{3(1 - e_o^2)^2}\right)\right\} \tag{28-40}$$

Thus $\frac{\Delta P}{P}$ measurements can be related to ρ_{r_p} and H. From two or more satellites with different perigee heights, or from one satellite at separated dates, one can determine density as a function of altitude.

$$\rho = \rho_{r_p} \exp\left[-\frac{1}{H}(r - r_p)\right] \tag{28-41}$$

If one includes atmospheric rotation and planetary flattening, the equivalent to (28-34) becomes

$$\frac{\Delta P}{P} = -\frac{3}{2} a_o B\delta(1-d)^2 \int_{-\pi}^{\pi} \rho(r) \frac{\left[1 + \frac{1+d}{1-d} e_o \cos E^{(o)}\right]^2}{\left[1 - e_o^2 \cos^2 E^{(o)}\right]^{1/2}} \, dE^{(o)} \tag{28-42}$$

with

$$d = \frac{\omega_e}{n_o}\sqrt{1-e_o^2}\ \cos i_o \; . \tag{28-43}$$

Again we can integrate to give $\frac{\Delta P}{P}$ as a function of $\frac{a_o e_o}{H}$, e_o, etc., as before, but it is a real mess. See Sterne's book, page 162, or his article in Journal of American Rocket Society, Vol. 29, page 777 (1959). A very readable account of these matters is given by G. E. Cook in "Effect of an Oblate Rotating Atmosphere on the Orientation of a Satellite Orbit," June 1960, RAE Tech Note GW 550.

The satellite lifetime is defined as the time when $e = 0$ where

$$e = e_o + \frac{\Delta e}{T} t_L = 0 \tag{28-44}$$

Repeating the previous analysis for $\frac{\Delta e}{T}$ and solving for t_L gives (for $\frac{a_o e_o}{H} > 2$).

$$t_L = -\frac{3 e_o P_o}{4\ \Delta P}\left[1 + \frac{7}{5}e_o + \frac{5}{16}e_o^2 + \frac{H}{2e_o a_o}\left[1 + \frac{11}{12}e_o + \frac{3H}{4a_o e_o}\right]\right] \tag{28-45}$$

and for $\frac{a_o e_o}{H} < 2$

$$t_L = -\frac{3\, e_o P_o I_o\left(\frac{a_o e_o}{H}\right)}{4\,\Delta P\, I_1\left(\frac{a_o e_o}{H}\right)} \tag{28-46}$$

$\frac{\Delta P}{P_o}$ is approximately 10^{-5} for low altitude satellites and 10^{-7} to 10^{-8} for higher ones. Therefore one must take care in obtaining these measured values. A good discussion of this point is made by Luigi Jacchia in "The Determination of Atmospheric Drag on Artificial Satellites," pages 136-142 of the book Dynamics of Satellites edited by M. Roy, Academic Press 1963.

In order to evaluate equations like (28-40), it is necessary to have a value for $B = \frac{C_D A}{m}$. To evaluate A one has to pick the effective cross sectional area of the satellite. If a long thin satellite rotates about a transverse axis, two possible modes are possible: (a) travelling like an airplane propeller and (b) tumbling end over end. For a cylinder of length ℓ and diameter d under mode (a), $A = \ell d$ and for mode (b) the effective cross sectional area A perpendicular to the direction of motion is $A = \frac{2}{\pi} (\ell d + \frac{1}{4} \pi d^2)$. For other possible modes, A lies between these extremes. One usually takes A as the mean value under motions (a) and (b).

$$A = \ell d \; [0.813 + \frac{d}{4\ell}] \quad .$$

The mass, m, of the satellite is assumed known.

The evaluation of the drag coefficient C_D is discussed in detail by Cook (see reference number 1). In the free-molecule flow regime which is appropriate for satellites with perigee heights greater than 200 km., the drag coefficient is influenced by five factors:

(1) The molecular speed ratio, i.e., the satellites speed divided by the most probable molecular speed. This is easily calculated.

(2) The mechanism of molecular reflection, specular or diffuse.

(3) The accommodation coefficient, which is the ratio of the change in the energy of molecules between striking the surface and re-emission, to the energy change they would suffer if emitted at the surface temperature.

(4) The temperature of the satellite.

(5) The fraction of molecules dissociated by impact.

After surveying the available evidence, Cook chose diffuse reflection, accommodation coefficient near 1, ignored any dissociation of molecules and assumed a temperature of 0°C, although the results are insensitive to this latter parameter. With these assumptions it is found that for spheres, for cones rotating about a transverse axis, and for a cylinder of ℓ/d between 5 and 20 spinning about a transverse axis, the appropriate value of C_D is between 2.1 and 2.3. Most analysis use $C_D = 2.2$.

With these constants one can then calculate B and find ρ_{r_p}. For most satellites B lies between 50 and 200.

If the air density were constant from day to day, the ΔP value would increase slowly and smoothly as the perigee height and eccentricity gradually decreased. In practice ΔP is found to be markedly irregular. Two major causes for these variations are now well documented: the air density is linked with solar activity and it varies between night and day.

There is a permanent thermal bulge in the atmosphere located in the bright hemisphere - toward the sun - which causes a day-night (durinal) variation in air density. The density peaks rather sharply around 2 PM solar time; the night time minimum is flatter than the maximum and occurs in the second half of the night, between 2 AM and 5 AM. The sun is directly overhead at 12 noon solar time. See Figure 28-3.

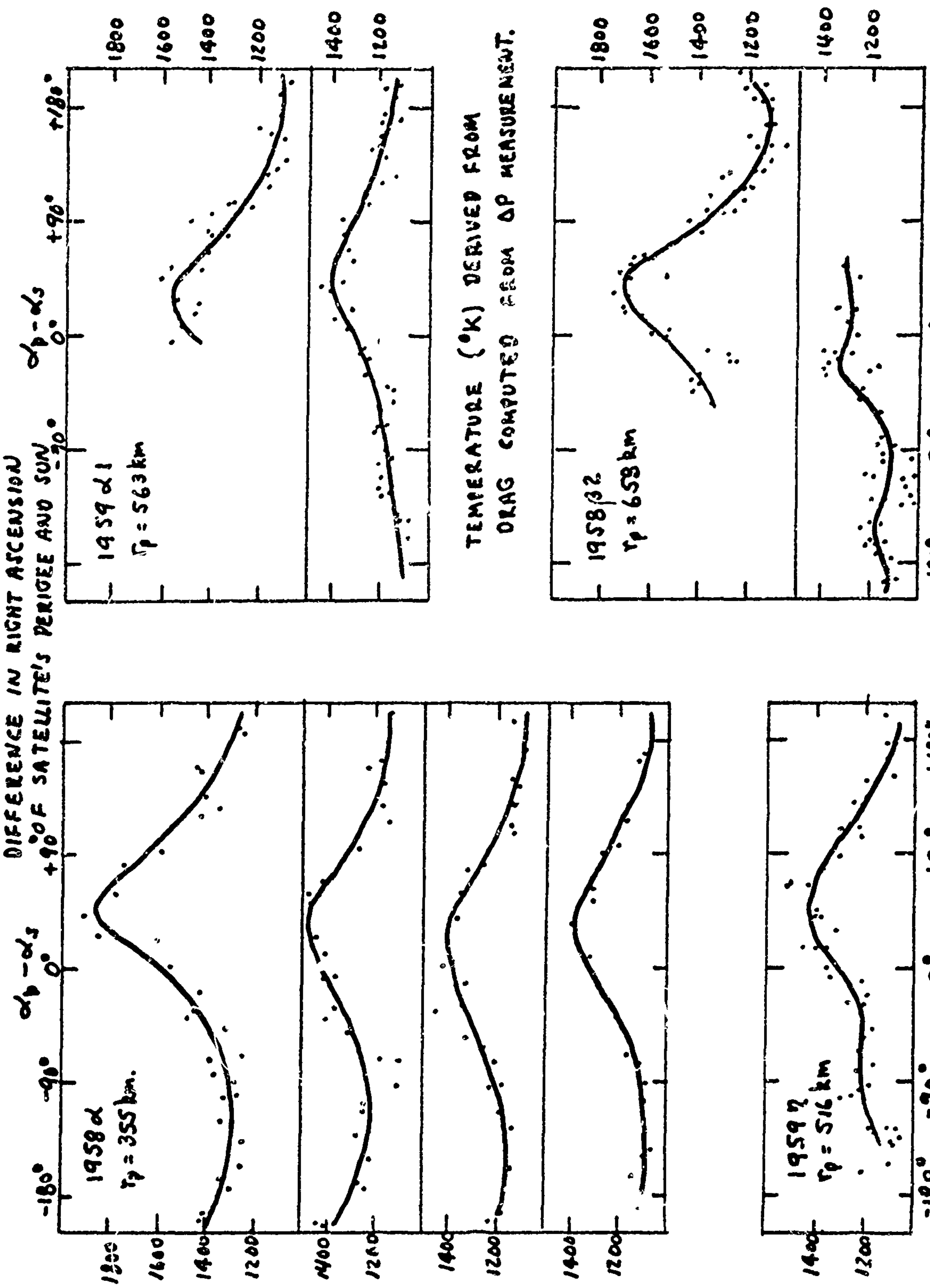


Figure 28-3

THE DIURNAL EFFECT FOR FOUR SATELLITES.

In the height region from 200 to 700 km., the atmospheric density at a given height y can be satisfactorily represented by

$$\rho = \rho_o \left[1 + f(y) \cos^6 \frac{\psi}{2} \right]$$

where ρ_o is the minimum night time density and ψ is the geocentric angular distance from the center of the durinal bulge. The temperature distribution appears to fit a curve of the form

$$T = T_o \left(1 + 0.4 \cos^4 \frac{\psi}{2} \right)$$

where T_o is the night time minimum temperature.

The bulge also shifts with the seasons making the density a function of the seasons as well. This is a latitude dependent effect.

The durinal effect depends on the height and is a direct consequence of conduction heating. Solar radiation primarily from solar lines λ304 (He II) and λ584 (He I) heat the atmosphere in roughly the F 1 layers in the region 100 to 200 kilometers above the earth. Below 180-200 km. the day-night effect is barely detectable. In the range 200-250 km. it is perhaps 10% higher in sunlight than dark. At 400 km the day-night difference is about 1.5 to 1, at 600 km it is 6:1 and at 700 km it is more like 10:1 or more. The solar radiation heats the F 1 layer and the atmosphere is heated by conduction following diffusion equilibrium and is assumed isothermal for any geographical location.

The Vanguard satellite has a perigee of about 658 kilometers. In October 1958 the perigee was at the center of the bulge and had a value of $\rho_{r_p} = 1.5 \times 10^{-15}$ g/cm^3. In June 1960 the perigee was in the dark

hemisphere and $\rho_{r_p} = 2.5 \times 10^{-17}$ g/cm^3. The possible density variations are more like 1000 to 1. This is due not only to the durinal bulge but to other solar radiation effects. 1958 was a year of high solar activity while June 1960 was more of a quiet year.

Erratic fluctuations in orbital accelerations of Sputnik II lead to the discovery of the correlation of drag with variation in solar radiation. A periodicity of about 27 days in the drag pointed to the same variation in solar radiation. The radiation level variations causes density fluctuations. See Figure 28-4.

A correlation was found between the satellite acceleration (and hence drag) and the solar flux measured at the wavelength of 20 cm. It was confirmed for the 10.7 cm flux and over an interval of 10 months all the individual maxima and minima of the solar flux curve had their counterpart, in phase, in the drag curve. See Figure 28-4.

In addition to this effect, Jacchia found two transient increases in drag coincided in time and duration with two violent magnetic storms. It now appears that all geomagnetic perturbations, even the smallest, effect the denisty of the upper atmosphere. This effect is shown in Figure 28-6. The effect on drag lags the increase in the a_p index by about five hours, where a_p is a measure of the geomagnetic index. At 600 kilometers, the air density can increase for a few hours by a factor of up to 8 during severe magnetic storms. These are not predictable.

There also seems to be a semi-annual variation in drag caused by solar winds, with a minima in January and July and a maximum in April

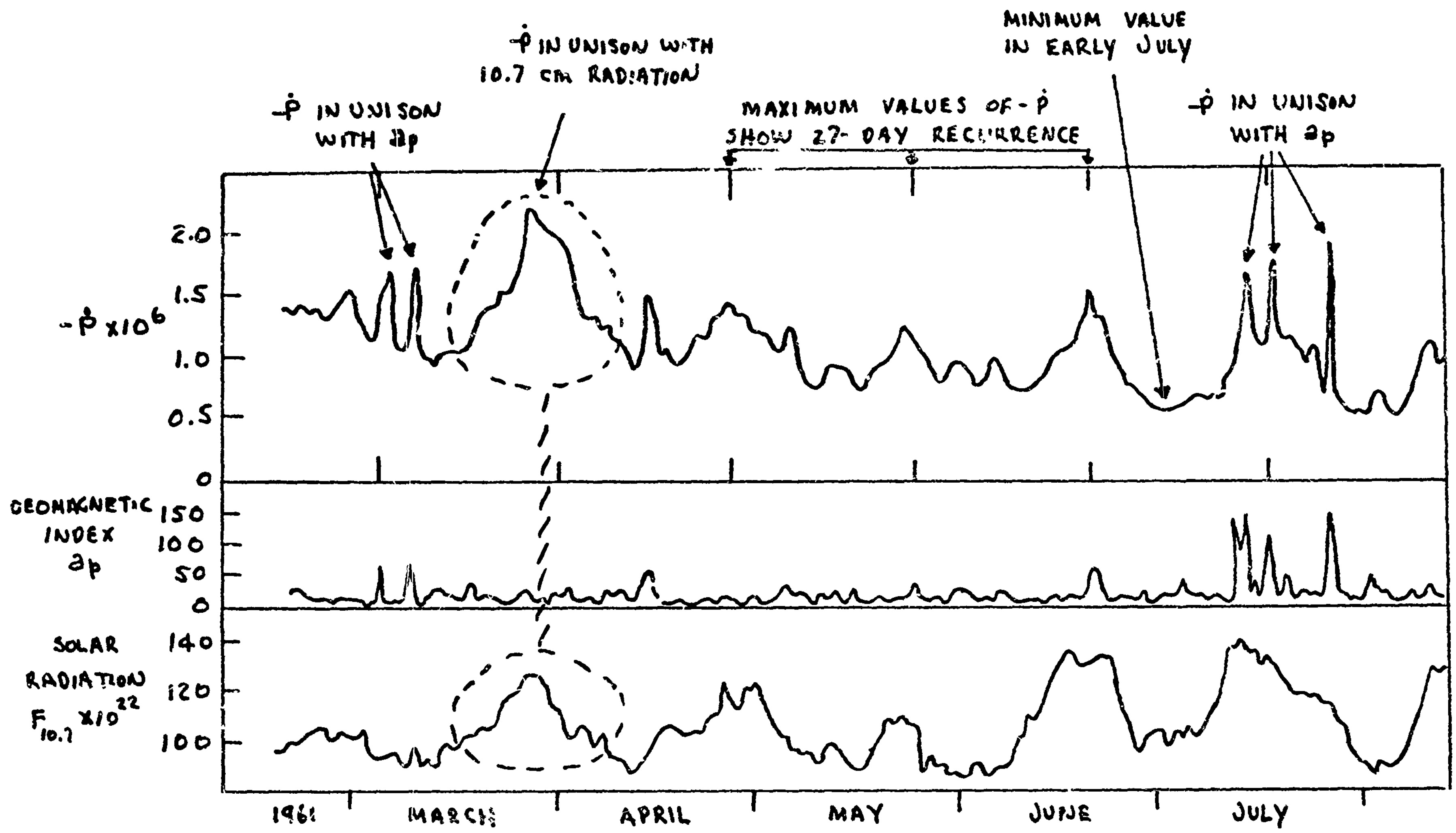
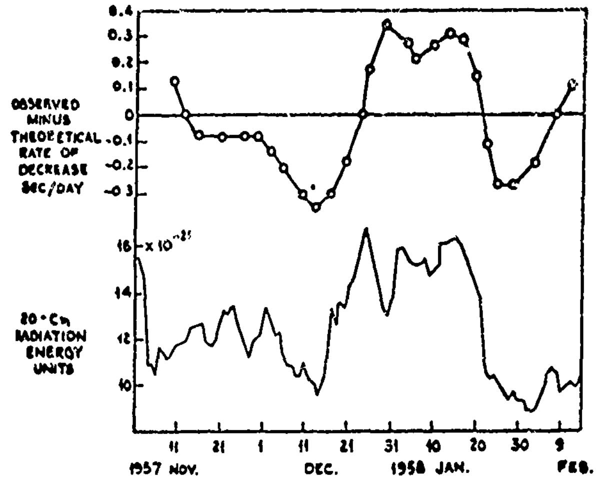
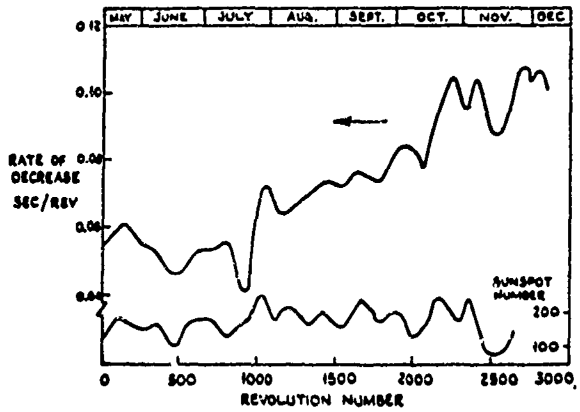


FIGURE 28-4 RATE OF CHANGE OF ORBIT PERIOD DUE TO AIR DRAG FOR EXPLORER 9



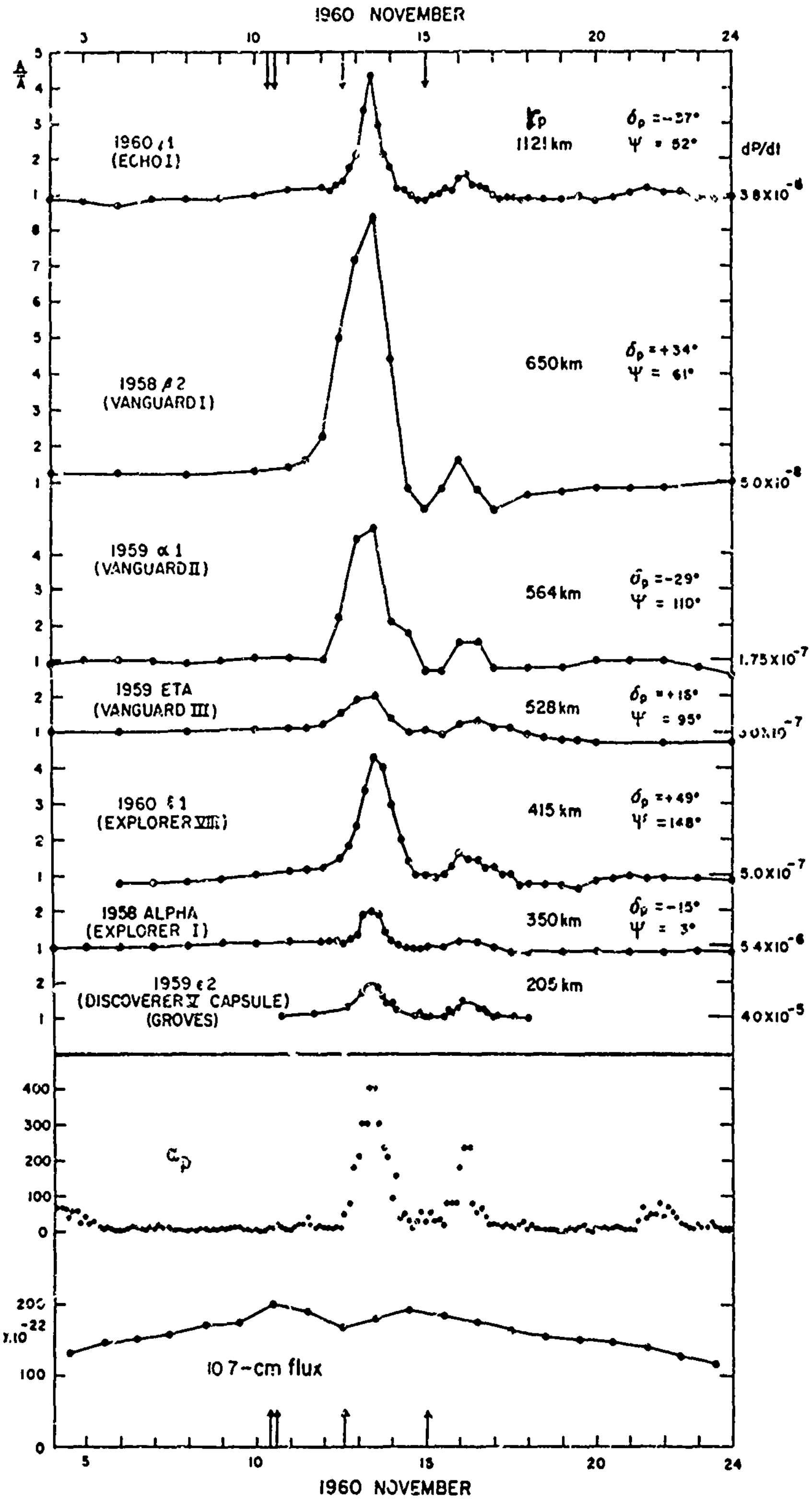
(a) Sputnik 2, and 20 cm solar radiation.
(After PRIESTER [1959].)



(b) Sputnik 3, and sunspot numbers.
(After PAETZOLD [1959a].)

Rate of decrease of satellite orbital period, compared with indices of solar activity.

FIGURE 28-5



Atmospheric drag of seven artificial satellites during the November 1960 events, compared with the geomagnetic planetary index a_p and the solar flux at 10.7 cm.

FIGURE 28-6

and October which are in phase with statistical variations of planetary geomagnetic index. See Figure 28-12.

To represent all of these variations we assume the empirical model that solar radiation heats the atmosphere at about 180 kilometers. The temperature source has the value

$$T = T_N \left(1 + 0.4 \cos^4 \frac{\psi}{2} \right) + 1.2 a_p$$

where a_p is the three hour geomagnetic index, ψ is the geocentric angular distance from the center of the durinal bulge, and T_N is the night time temperature given by

$$T_N = 330^\circ + 2.1 F_{10.7} + \left(2.0 + 0.5 \cos 2\pi \frac{t - \text{April } 7}{365} \right) \bar{F}_{10.7}$$

where

$F_{10.7}$ is the daily 10.7 cm solar flux in units of 10^{-22} watts/m^2/cycle/second. This is the 27 day fluctuation.

$\bar{F}_{10.7}$ is the monthly average and represents the solar wind effect.

The night time temperature can vary from 700°K during solar minimum activity to 1500°K during intense sunspot activity. The maximum daytime temperatures are about 35% higher.

The important point of this discussion is that no static model can hope to represent the density variations. Thus when we choose H as a constant we are in trouble. Figure 28-8 shows variations of H with

1. K.S.W. Champion and R. A. Minzner, "Revision of United States Standard Atmosphere, 90 to 700 Kilometers," Rev. Geophys., 1 (Feb. 1963), 57-84.
2. V. I. Krassovsky, "Some Geophysical and Astronomical Aspects of Soviet Space Research," Endeavour, April 1962, p. 67.
3. K. Moe, "A Model for the Errors in Orbital Predictions Caused by Fluctuations in Drag," Nature, 192 (Oct. 14, 1961), 151.
4. N. W. Spencer, "Aeronomy Research with Rockets and Satellites," Proc., NASA-University Conference on the Science and Technology of Space Exploration, Nov. 1 ' 1962, Vol. 1, pp. 147-151.
5. J. E. Sterne, "Effect of the Rotation of a Planetary Atmosphere Upon the Orbit of a Close Satellite," ARS J., 29 (1959), 777.

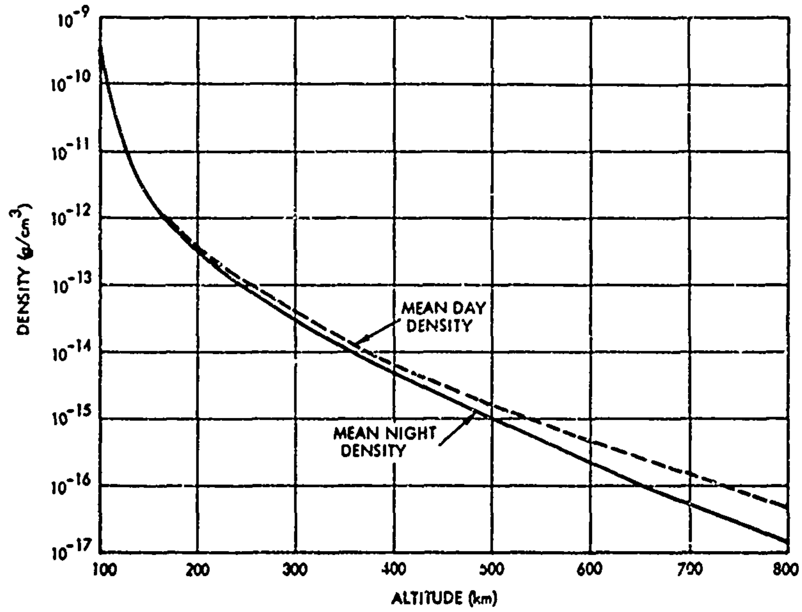


Fig. A Atmospheric density as a function of altitude.

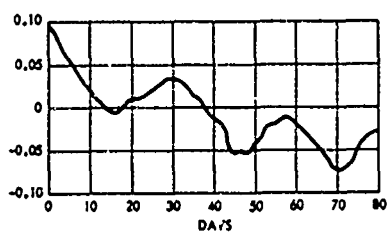


Fig. B Autocorrelation function of rate of change of drag observed on Explorer 9.

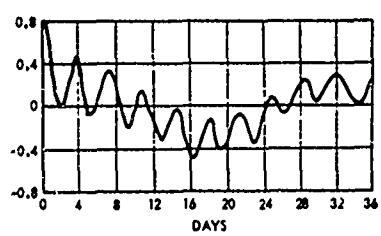


Fig. C Autocorrelation function of rate of change of drag observed on Explorer 1.

FIGURE 28-7

height, with day-night, and with enhanced solar activity. Figure 28-9 shows some early density profiles as determined by various authors and Figure 28-10 shows the corresponding variations of H for these density profiles.

King-Hele has revised the analysis of this section to include the linear variation of H with altitude in order to improve the analytical results. See his article, "The Contraction of Satellite Orbits under the Action of Air Drag, Allowing for the Variation of Scale Height with Altitude," pages 211-218 of the book Dynamics of Satellites edited by M. Roy, Academic Press, 1963.

King-Hele also demonstrated that a more accurate determination of the density could be made by evaluating it not at perigee, r_p, but at a point $r_p + \frac{H}{2}$ in height, where a small change in the calculated value of H would cause only a very small in density determination. The reason, which is not immediately obvious, is that this point is close to the weighted mean of the heights over which drag is effective. Figure 28-11 shows that the drag acts primarily about the perigee point.

If we let H* be the best estimate of H, then the density ρ_λ at a height of $r_p + \lambda H^*$ is

$$\rho_\lambda = \rho_{r_p} \exp\left[- \frac{\lambda H^*}{H} \right]$$

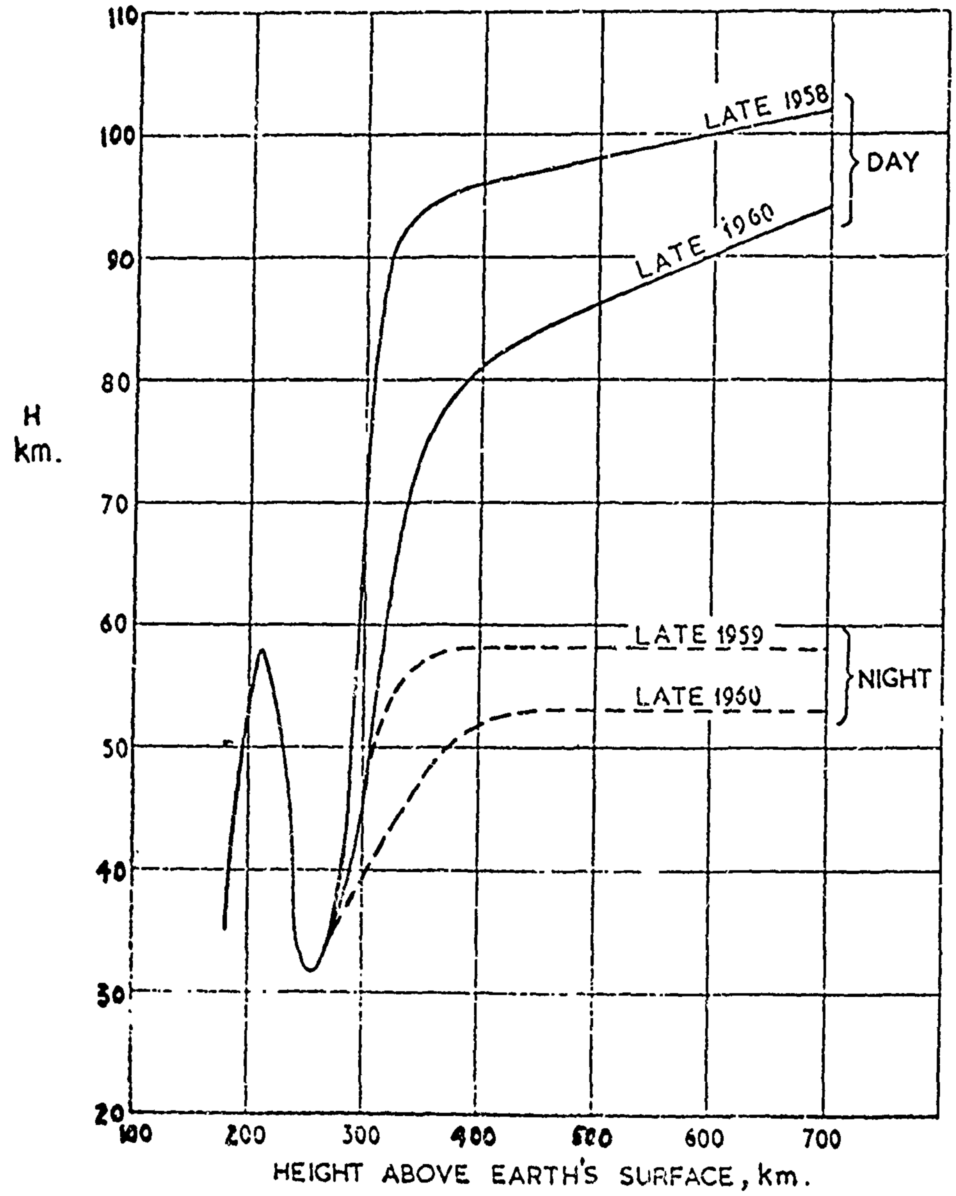
From this King-Hele shows that one can calculate the following:

$$\rho_{r_p} = - \frac{\Delta P}{3 P_0 B \delta} \sqrt{\frac{2 e_0}{\pi a_0 H^*}} \left[\frac{H^*}{H} e^{-\frac{\lambda H^*}{H}} \right] \left\{ 1 + 2 e_0 - \frac{H}{8 a_0 e_0} + O\left(e_0^2, \frac{H^2}{a_0^2 e_0^2} \right) \right\}$$

For $\lambda = \frac{1}{2}$ the term in the brackets [ ] shows very little change when $\frac{H^*}{H}$ varies between 0.6 and 1.5 so we can replace H by H* to give

$$\rho_{r_p} = - \frac{0.158 \Delta P}{P_0 B \sigma} \sqrt{\frac{e_0}{a_0 H^*}} \left[1 + 2e_0 - \frac{H^*}{8a_0 e_0} + O\left\{ e_0^2, \left(\frac{H^*}{a_0 e_0}\right)^2 \right\} \right]$$

Solar radiation pressure affects also directly effect the orbital period and must be carefully separated from the drag effects. For a relatively close satellite with moderately eccentric orbit ($0.1 < e < 0.2$) the variations in ΔP caused by solar radiation pressure on the satellite are of the order of $\pm 1 \times 10^{-7} \frac{A}{m}$ when $\frac{A}{m}$ is in cm^2/g. For comparison, the atmospheric drag at intermediate heights has ΔP of the order of $\pm 1 \times 10^9 \rho \frac{A}{m}$, where ρ is the atmospheric density in g/cm^3. Thus for ρ of the order of 10^{-16} g/cm^3 the solar radiation pressure effects may equal that of atmospheric drag. At times of sunspot maximum this will occur at about 900 kilometers; at low solar activity, when the atmosphere is appreciably contracted, it will occur as low as 500 kilometers above the earth. To compute the drag to $\pm$ 10% or better, one must account for solar radiation pressure effects whenever the perigee height is greater than 400 kilometers.



VARIATION OF H WITH HEIGHT

FIGURE 28-8

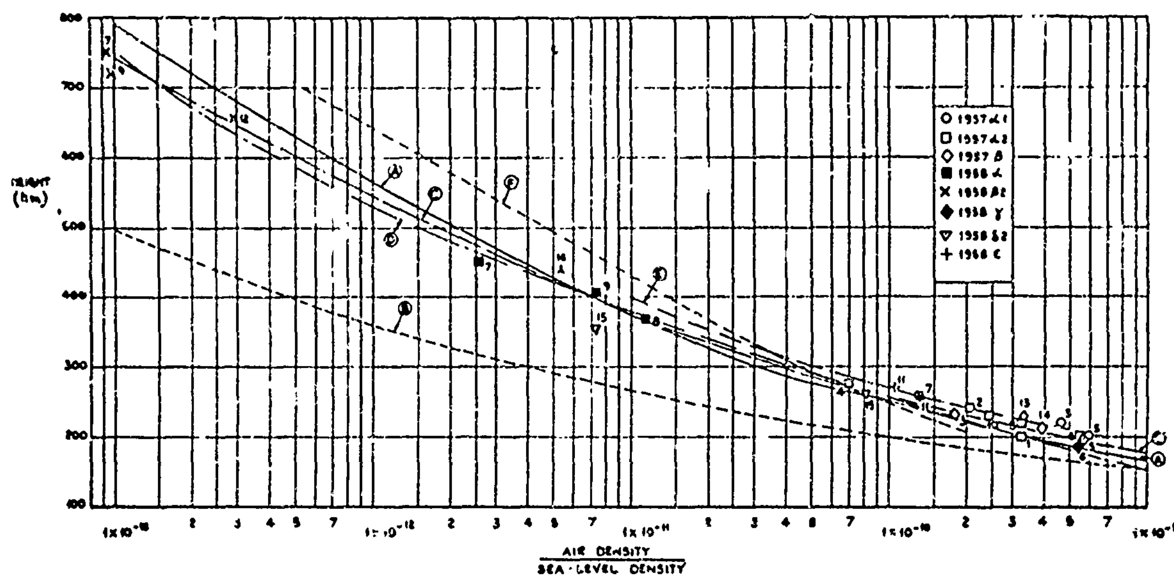


Fig. 3.2. Values of air density obtained by means of satellites, with proposed "standard atmospheres". (For key to numbers and letters, see Table 3.3

TABLE 3.2

Values of air density and H given by the curves of Fig. 3.1

Height (km)	Density (g/cm^3)			H (km)		
180	7.0×10^{-13}			30		
190	5.1×10^{-13}			36		
200	3.9×10^{-13}			43		
210	3.2×10^{-13}			47		
220	2.5×10^{-13}			44		
230	2.0×10^{-13}			42		
240	1.5×10^{-13}			41		
250	1.2×10^{-13}			40		
260	9.4×10^{-14}			41		
270	7.2×10^{-14}			42		
280	5.7×10^{-14}			44		
290	4.6×10^{-14}			47		
300	3.7×10^{-14}			50		
	day (early 1959)	day (mid 1960)	night (1959–1960)	day (early 1959)	day (mid 1960)	night (1959–1960)
300	3.7×10^{-14}	3.7×10^{-14}	3.7×10^{-14}	50	50	50
350	1.7×10^{-14}	1.6×10^{-14}	1.4×10^{-14}	73	66	54
400	9.2×10^{-15}	8.0×10^{-15}	5.5×10^{-15}	84	73	↓
450	5.1×10^{-15}	4.0×10^{-15}	2.1×10^{-15}	93	77	
500	3.1×10^{-15}	2.1×10^{-15}	8.1×10^{-16}	101	81	
550	1.9×10^{-15}	1.1×10^{-15}	3.1×10^{-16}	109	86	
600	1.2×10^{-15}	6.7×10^{-16}	1.2×10^{-16}	118	94	
650	8.1×10^{-16}	4.0×10^{-16}	4.7×10^{-17}	129	106	
700	5.6×10^{-16}	2.6×10^{-16}	1.8×10^{-17}	143	123	54

Sea-level density taken as 0.001 226 g/cm^3.

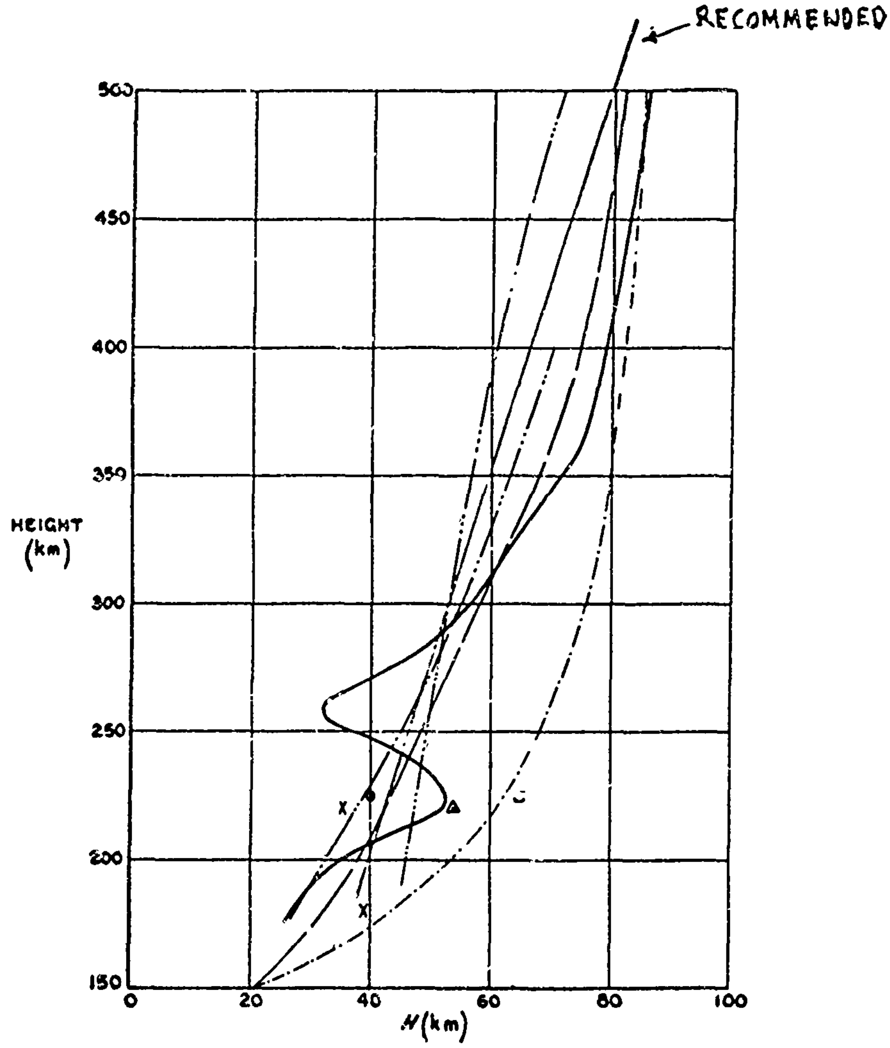
TABLE 3.3

Key to number and letters in Fig. 3.2

No. in Fig. 3.2	Authors
1	Mullard Observatory [1957]
2	Royal Aircraft Establishment [1957]
3	STERNE and SCHILLING [1957]
4	HARRIS and JASTROW [1958a]
5	GROVES [1958a]
6	STERNE [1958c]
7	SIRY [1958]
8	STERNE [1958b]
9	HARRIS and JASTROW [1958b]
10	SEDOV [1958]; LIDOV [1958]
11	SCHILLING and WHITNEY [1959]
12	JACCHIA [1958b]
13	ELYASBERG [1958]
14	SCHILLING and STERNE [1959]
15	KRASSOVSKY [1959]; MICHNEVICH [1958]
16	KURT [1959]

Letter in Fig. 3.2	Authors
A	KING-HELE [1959c]
B	MINZNER and RIPLEY [1956]
C	KALLMANN [1959]
D	GROVES [1959a]
E	SCHILLING and STERNE [1959]
F	HARRIS and JASTROW [1959]

FIGURE 28-9



Key: ——··—— SCHILLING and STERNE [1959]
—·—·—· HARRIS and JASTROW [1959]
——···—— GROVES [1959a]
—— —— KALLMANN [1959]
———— KING-HELE [1959c]
⊙ LIDOV [1958]
⊡ KING-HELE and LESLIE [1958]
△ NONWEILER [1959]
× GROVES [1959b]

Variation of H with height, as given by various authors.

FIGURE 28-10 (see Figure 28-9)

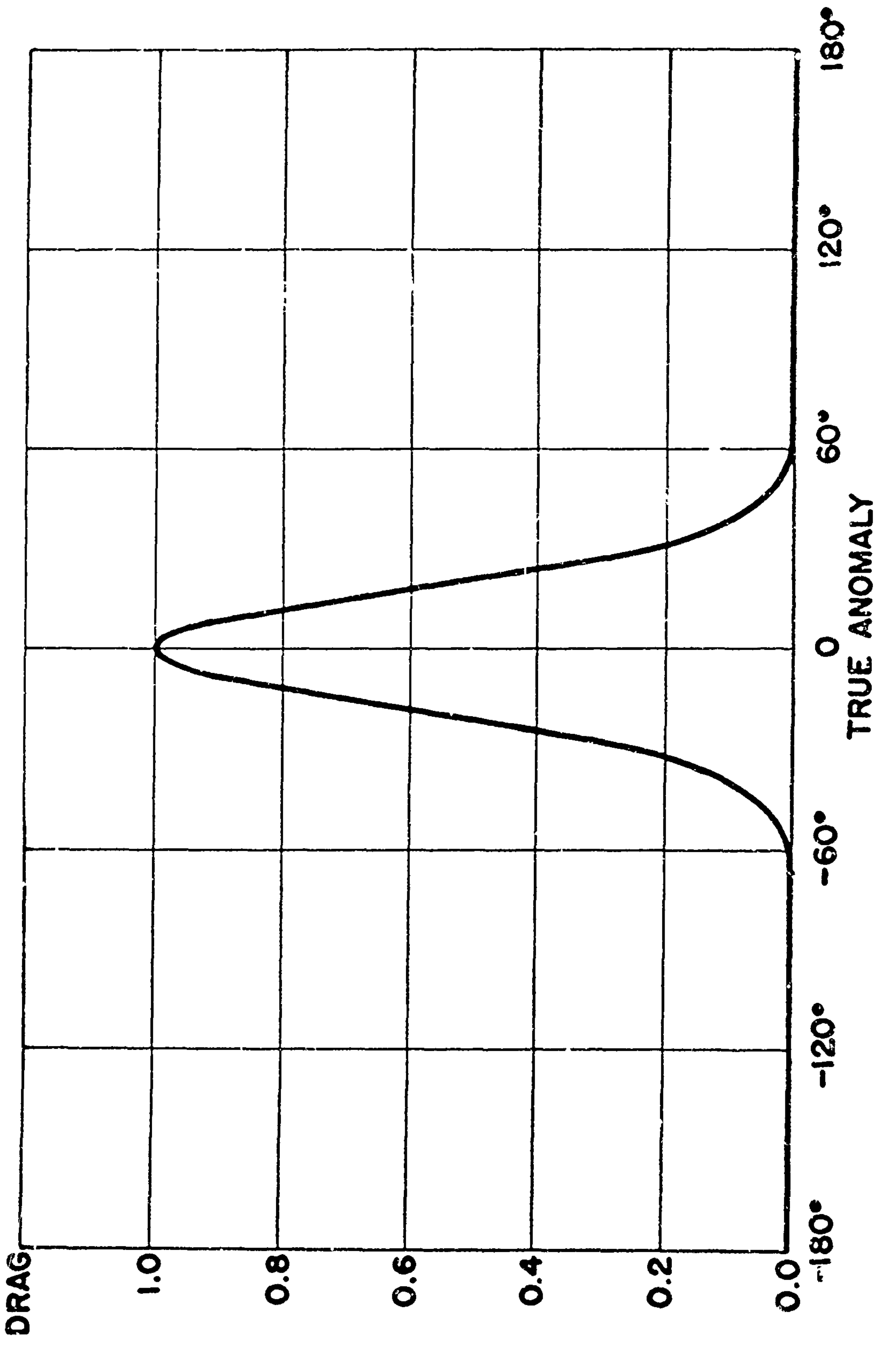
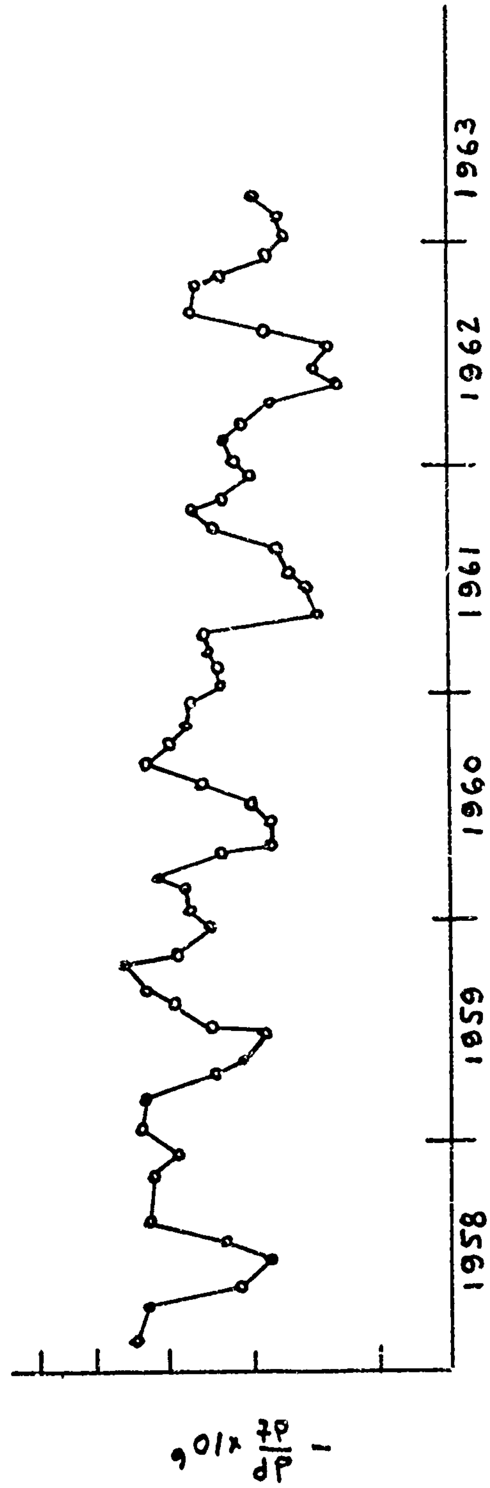


FIGURE 28-11



RATE OF CHANGE OF ORBITAL PERIOD FOR EXPLORER I. CORRECTED TO ALLOW FOR VARIATIONS IN SOLAR ACTIVITY, GEOMAGNETIC INDEX AND DAY-TO-NIGHT EFFECT. (PAETZOLD)

FIGURE 28-12

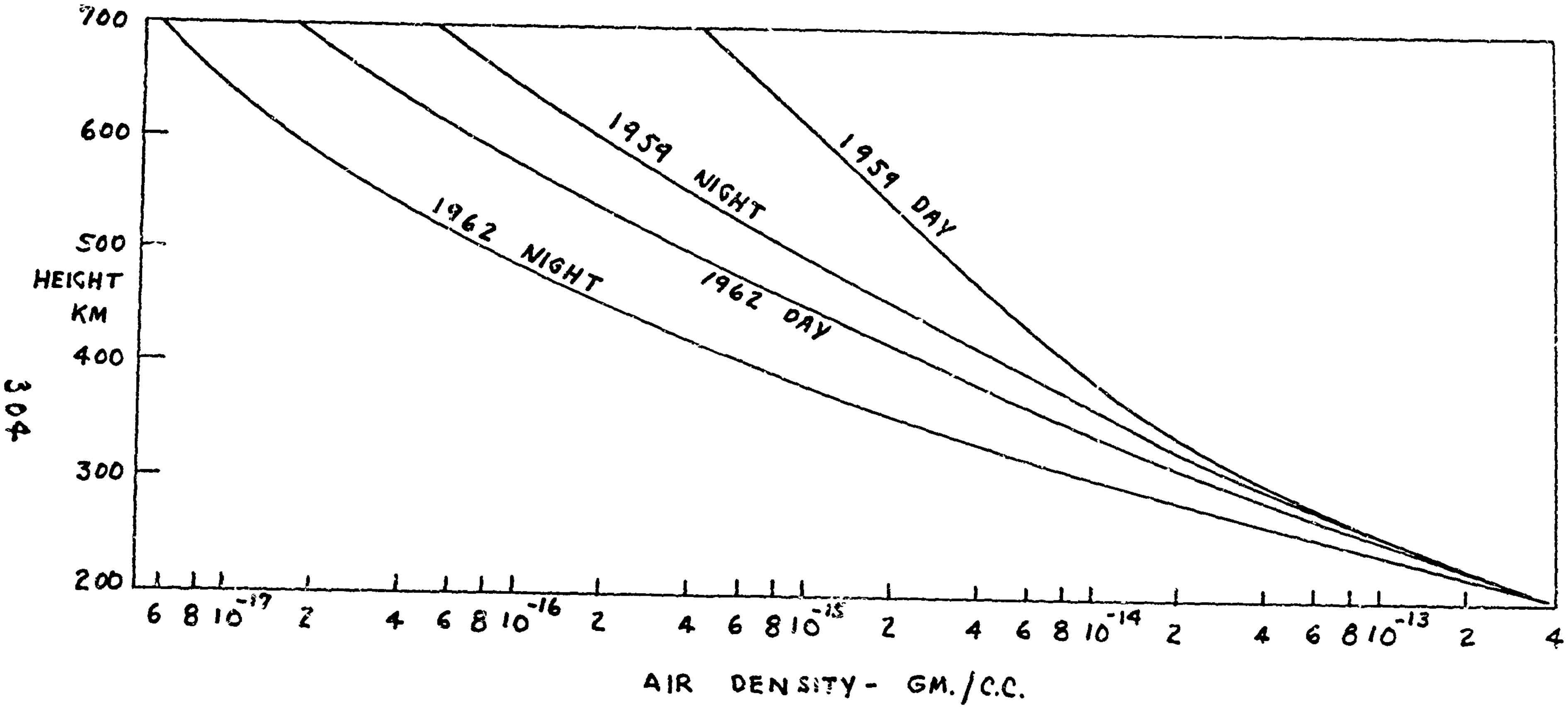


FIGURE 2B-13 DAY-TO-NIGHT VARIATION IN DENSITY FOR 1959 AND 1962. (KING-HELE)

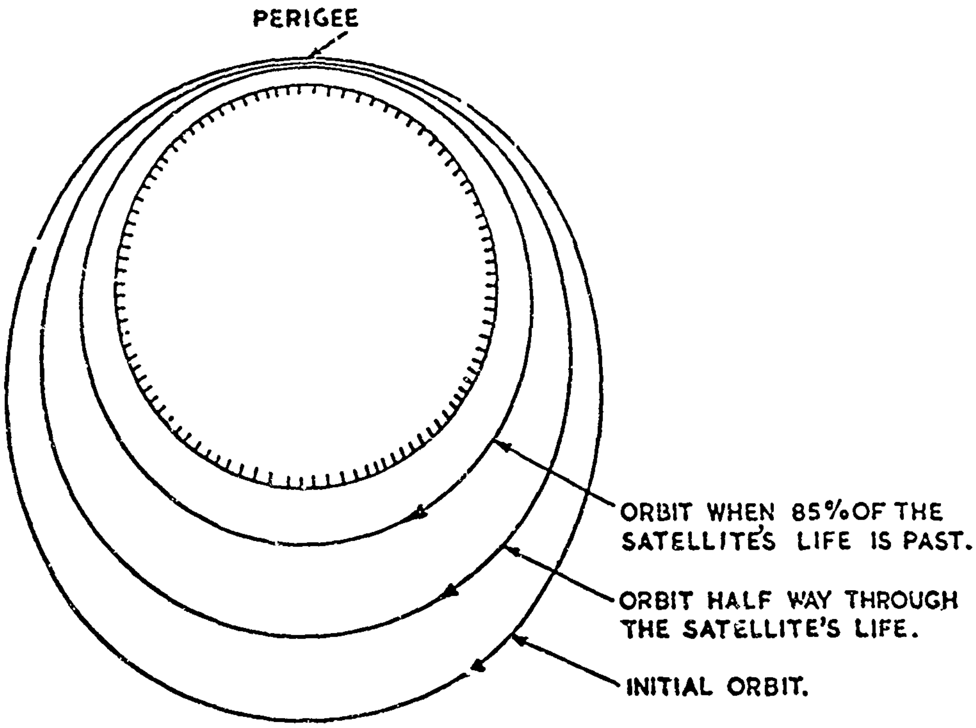


FIG. 28-14 CONTRACTION OF SATELLITE ORBIT UNDER THE ACTION OF AIR DRAG

REFERENCES

1. Cook, C. E., "Satellite Drag Coefficients," Royal Aircraft Establishment (RAE) Technical Report No. 65005, January 1965, AD 464391.

2. King-Hel , D. C. and C. E. Cook, "The Contraction of Satellite Orbits under the Influence of Air Drag Part IV with Scale Height Dependent on Altitude," RAE Technical Note No. Space 18, September 1962.

3. Jacchia, L. G., "Variations in the Earth's Upper Atmosphere as Revealed by Satellite Drag," Smithsonian Astrophysical Observatory Report, December 31, 1962, AD 296528.

4. King-Hele, D. G., "The Upper Atmosphere - A Review," RAE Technical Note Space 50, October 1963, AD 428871.

5. King-Hele, D. C., "Improved Formulae for Determining Upper Atmosphere Density from the Change in a Satellite's Orbital Period," RAE Technical Note Space 21, 1962. See also Planet.Space Science, Vol. 11, page 261 (1963).

6. Harris, I. and W. Priester, "Theoretical Models for the Solar Cycle Variations of the Upper Atmosphere," NASA Technical Note D-1444 (1962).

7. Paetzold, H. K., "New Results about the Annual and Semi-Annual Variation of the Upper Terrestrial Atmosphere," paper pr sented at 4th International Space Science Symposium, Warsaw, 1963.

29. VON ZEIPEL'S METHOD IN PERTURBATION THEORY

The von Zeipel method consists of making successive mathematical transformations of variables of a canonical system in a methodical way so that the final solution is obtained in a certain desired form. In particular one wishes to separate secular terms from very long period terms in the answer. These latter periodic effects of say 100 year periods would be difficult to separate from the secular effect by observational means.

In the analytical investigations of the oblateness effect, certain elements such as ω, Ω, and χ experienced secular variations plus periodic fluctuations about these secular variations. Other elements had only periodic variations. Furthermore, there was a distinction among the periodic variations as to rapid or short period variations and slow or long period variations. To visualize these effects, consider Figure 29-1. Besides the secular term, there is a long-period variation caused by the continuous variation of ω, since the elements vary as trigonometric functions of ω and these can have very long periods for certain inclination angles. In addition there are short period variations on top of the long period ones which are caused by trigonometric functions which are linear combinations of M or f and ω. These fast variations are caused by variations in the true anomaly which are much more rapid than the slow secular variations in ω. We thus have the solution in the form of

$$q = q_o + \dot{q}_o (t-t_o) + K_1 \cos 2\omega + K_2 \sin (2f+2\omega) + \dots \qquad (29\text{-}1)$$

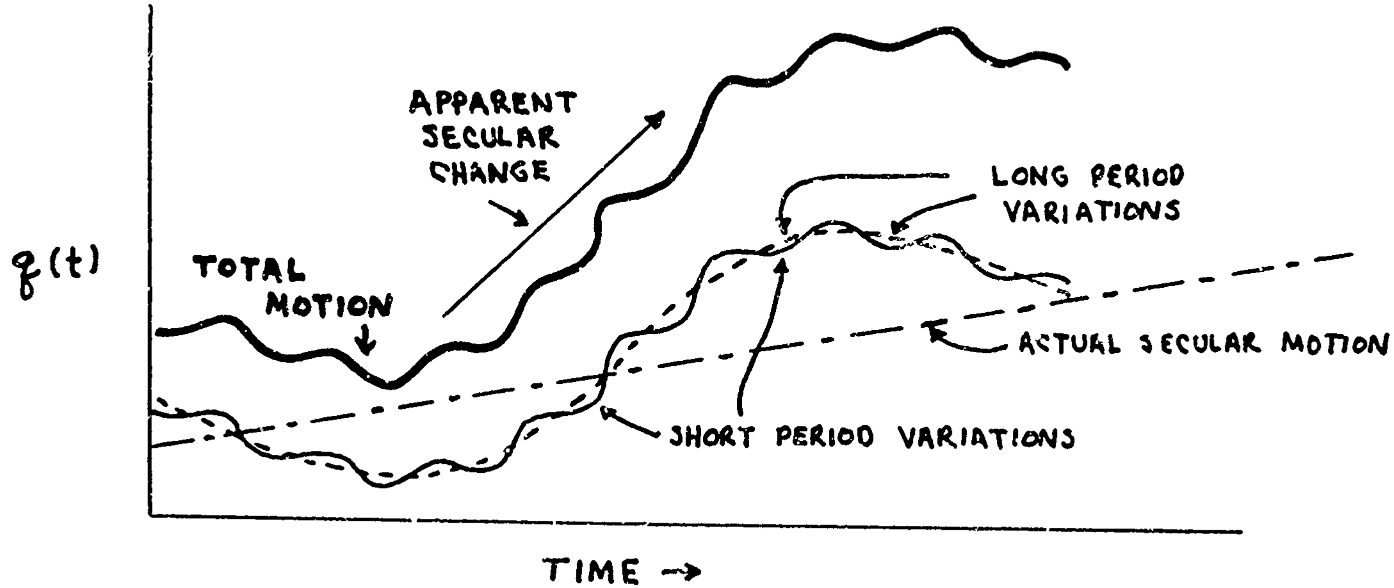

g(t)
TOTAL MOTION
APPARENT SECULAR CHANGE
LONG PERIOD VARIATIONS
ACTUAL SECULAR MOTION
SHORT PERIOD VARIATIONS
TIME →

FIGURE 29-1

The first term is the adopted epoch mean value, the second term is the secular variation, the third term is the long period variation and the rest are presumably short period variations. In the von Zeipel method we are able to separate the secular terms from the periodic variations and then separate the long and short period terms.

Let's assume we have a system with two degrees of freedom which is represented by only two canonical equations

$$\frac{d\ell}{dt} = \frac{\partial H}{\partial L} \qquad \text{and} \qquad \frac{dL}{dt} = - \frac{\partial H}{\partial \ell} \tag{29-2}$$

where $H = H\,(L,\ell)$ is the system Hamiltonian.

We can express this as

$$H(L,\ell) = H_o(L) + [H(L,\ell) - H_o(L)]. \tag{29-3}$$

The $H_o(L)$ represents some unperturbed portion with the important feature that it contain only one of the canonical variables.

The first step in the von Zeipel method as modified by Brouwer, is to transform the Hamiltonian H into a new Hamiltonian H* so that one or more of the state variables present in H are eliminated from H*. In the case of several canonical variables the operation may progress in stages by transforming H* into H**, etc., until all momenta state variables are present in the final result only in the form of constants.

For example, $H(L, \ell)$ would be transformed into $H^*\,(L', -)$, where the dash indicates the absence of ℓ'. We are transforming from the old canonical system with coordinates (L, ℓ), to a new canonical system with

coordinates (L', ℓ') in such a way that the new Hamiltonian H^* is a function of L' alone. Then because the transformation is a canonical one, the new Hamilton-Jacobi canonical equations are

$$\frac{dL'}{dt} = -\frac{\partial H^*(L',-)}{\partial \ell'} = 0 \tag{29-4}$$

and therefore

$$\frac{d\ell'}{dt} = \frac{\partial H^*}{\partial L'} = f(L') = \text{constant}. \tag{29-5}$$

From this we see that L' is a constant and ℓ' is of the form $\ell' = k\,(t-t_o)$. However we can't always find a generating function to do this so we proceed as follows:

Assume that $H = H_o + (H-H_o)$ may be expanded in an infinite series in powers of some small parameter ε. Each term of such a convergent expansion will have a coefficient involving some power of ε,

$$H = H_o + H_1 + H_2 + H_3 + \ldots. \tag{29-6}$$

where H_n is of order ε^n, i.e., the subscript refers to the power of ε involved in the coefficient. For example $H_2 = \varepsilon^2\hat{H}$ is designated as being of second order in ε. Note that $(H-H_o)$ is of order ε because H_o and H are presumed to be close, i.e., with ε of each other.

To transform H into H^* we use a generating function S and likewise assume $S\,(L', \ell)$ is developable in a Taylor's series in the neighborhood of $\varepsilon = 0$.

$$S = S(L'. \ell) = S_o + S_1 + S_2 + S_3 + \ldots. \tag{29-7}$$

where again the subscript refers to the power of ε involved in the coefficient of each term. S_n for $n > 0$ contains only periodic terms and each is of order ε^n.

In order to have a canonical transformation, the old variables (L, ℓ) must be related to the new variables (L', ℓ') by the relations $[S = S(L', \ell)]$.

$$L = \frac{\partial S}{\partial \ell} \tag{29-8}$$

$$\ell' = \frac{\partial S}{\partial L'} \tag{29-9}$$

Since we want $\ell = \ell'$ when $\varepsilon = 0$, we arbitrarily define $S_o = L'\ell$ and have

$$S = L'\ell + S_1 + S_2 + S_3 + \ldots. \tag{29-10}$$

$$L = L' + \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \frac{\partial S_3}{\partial \ell} + \ldots. \tag{29-11}$$

$$\ell' = \ell + \frac{\partial S_1}{\partial L'} + \frac{\partial S_2}{\partial L'} + \frac{\partial S_3}{\partial L'} + \ldots. \tag{29-12}$$

Thus the new and old corresponding variables differ by a quantity at least of the first order.

$$\begin{aligned} L - L' &= \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots. = O(\varepsilon) \\ \ell' - \ell &= \frac{\partial S_1}{\partial L'} + \frac{\partial S_2}{\partial L'} + \ldots. = O(\varepsilon). \end{aligned} \tag{29-13}$$

We require that S_n (L', ℓ) depend only on ℓ through trigonometric functions for any n > o. This avoids secular perturbations in the momenta L, e.g.,

$$L - L' = \frac{\partial(S-S_o)}{\partial \ell} \tag{29-14}$$

are then periodic functions of ℓ. Thus our solution will be of the form

$$\begin{aligned} L &= L' + \text{trigonometric terms periodic in } \ell \\ \ell &= \dot{\ } (t-t_c) + \text{trigonometric terms periodic in } \ell \end{aligned} \tag{29-15}$$

The transformation from (L, ℓ) into (L', ℓ') is a canonical one and since S does not contain time explicitly, ($\frac{\partial S}{\partial t} = 0$), the new Hamiltonian will not be changed in form, i.e.,

$$H(L,\ell) = H_o(L) - [H(L,\ell) - H_o(L)] = H^*(L',-). \tag{29-16}$$

We again assume that the new Hamiltonian H* can be developed in a Taylor's series in the neighborhood of ε = 0. Since the terms in the bracket of equation (29-16) are of order ε we can also write this as

$$H(L,\ell) = H_o(L) + H_1(L,\ell) = H_o^*(L') + H_1^*(L') + H_2^*(L') + \dots . \tag{29-17}$$

Upon substitution of L from equation (29-11) we can write the left side of equation (29-17) as

$$H_o(L' + \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \dots .) + H_1(L' + \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \dots, \ell). \tag{29-18}$$

Our next step is to expand each of these in a Taylor's series about the point (L', ℓ) which is within ε of the point (L, ℓ).

Recall for $u = x + h$, $v = y + k$ we have

$$f(x+h,\ y+k) = f(u,v) = f(x,y) + \ h \frac{\partial f(u,v)}{\partial u}\bigg|_{\substack{u=x\\v=y}} \ k \frac{\partial f(u,v)}{\partial v}\bigg|_{\substack{u=x\\v=y}} \tag{29-19}$$

$$+ \frac{1}{2!}\left[h^2 \frac{\partial^2 f(u,v)}{\partial}\bigg|_{\substack{u=x\\v=y}} + 2hk \frac{\partial^2 f(u,v)}{\partial u \partial v}\bigg|_{\substack{u=x\\v=y}} + k^2 \frac{\partial^2 f(u,v)}{\partial v^2}\bigg|_{\substack{u=x\\v=y}}\right] + \ldots.$$

To simplify the notation we can write this as

$$f(u,v) = f(x,y) + [h \frac{\partial f(x,y)}{\partial u} + k \frac{\partial f(x,v)}{\partial v}]$$

$$+ \frac{1}{2!}[h^2 \frac{\partial^2 f(x,y)}{\partial u^2} + 2hk \frac{\partial^2 f(x,y)}{\partial u \partial v} + k^2 \frac{\partial^2 f(x,y)}{\partial v^2}] + \ldots.$$

For the H_o term we have

$$x = L' \qquad h = \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots \qquad v = 0$$

$$y = k = 0 \qquad u = L' + \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots = L \tag{29-21}$$

For the H_1 term we have

$$x = L' \qquad h = \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots \qquad u = L' + \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots. = L$$

$$y = \ell \qquad k = 0 \qquad v = \ell \tag{29-22}$$

Carrying out the expansion of each term in (29-18) gives the following

$$H_o(L',-) + \left(\frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots\right) \frac{\partial H_o(L',-)}{\partial L} + \frac{1}{2}\left(\frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots\right)^2 \frac{\partial^2 H_o(L',-)}{\partial L^2}$$

$$+ \ldots . + H_1(L',\ell) + \left(\frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots\right) \frac{H_1(L',\ell)}{\partial L} + \tag{29-23}$$

$$+ \frac{1}{2}\left(\frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots\right)^2 \frac{\partial^2 H_1(L',\ell)}{\partial L^2} + \ldots$$

We then substitute this into equation (29-17) to obtain

$$H_o(L') + \frac{\partial S_1}{\partial \ell}\frac{\partial H_o(L')}{\partial L} + H_1(L',\ell) + \frac{\partial S_2}{\partial \ell}\frac{\partial H_o(L')}{\partial L} + \frac{1}{2}\left(\frac{\partial S_1}{\partial \ell}\right)^2 \frac{\partial^2 H_o(L')}{\partial L^2}$$

$$+ \left(\frac{\partial S_1}{\partial \ell}\right)\frac{\partial H_1(L',\ell)}{\partial L} + \ldots\ O(\varepsilon^3) = H_o^*(L') + H_1^*(L') + H_2^*(L') + \ldots\ O(\varepsilon^3) \tag{29-24}$$

By collecting terms of corresponding order of ε on both sides of equation (29-24) we obtain the following sets of equations:

$$H_o(L') = H_o^*(L') \tag{29-25}$$

$$\left(\frac{\partial S_1}{\partial \ell}\right)\frac{\partial H_o(L')}{\partial L} + H_1(L',\ell) = H_1^*(L') \tag{29-26}$$

$$\frac{\partial S_2}{\partial \ell}\frac{\partial H_o(L')}{\partial L} + \frac{1}{2}\left(\frac{\partial S_1}{\partial \ell}\right)^2 \frac{\partial^2 H_o(L')}{\partial L^2} + \left(\frac{\partial S_1}{\partial \ell}\right)\frac{\partial H_1(L',\ell)}{\partial L} = H_2^*(L') \tag{29-27}$$

etc.

This expansion may be continued indefinitely. The expression for the general differential equations of this von Zeipel's method is given to any order by a general term developed by Giacaglia in NASA report N64-28079 (TMX-55058) entitled, "Notes on von Zeipels method," June 1964. In the meantime the student should carry out the details to obtain the next two higher order equations.

The object now is to solve these resulting partial differential equations (which for our two variable examples is actually a set of ordinary differential equations) for S_1, S_2, etc., depending on the order of accuracy desired. When this is done the solution follows from

$$\begin{aligned} L &= L' + \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots \\ \ell' &= \ell + \frac{\partial S_1}{\partial L'} + \frac{\partial S_2}{\partial L'} + \ldots \end{aligned} \qquad (29\text{-}28)$$

Note that the equations in S_1, S_2, etc., are first order linear equations. S_1 is determined from (29-26) and then substituted into (29-27) which is then first order and linear in S_2. Similarly the next set of equations (not shown here) can be solved using the determined values of S_1 and S_2. If we separate H_1 (L', ℓ) into a secular term and sum of periodic terms we can always solve the resultant partial differential equation by the method of characteristics. Further one can prove that the resultant series expansion for the variables, equation (29-28) converges. See, "A Proof of the Convergence of the Poincare' - von Zeipel Procedure in Celestial Mechanics" by Richard Barrar (January 1965), System Development Corporation Report SP-1926, AD-610694.

It might be well to again call attention to the notation used in the series terms. Terms like $\frac{\partial H_1(L',\ell)}{\partial L}$ are actually shorthand notation for $\frac{\partial H_1(L,\ell)}{\partial L}$ evaluated at $(L = L', \ell = \ell)$. Thus given H_1 (L, ℓ) one takes the partial with respect to L and then evaluates at the limit as $h \to 0$, i.e. $L = L' + h$ with $h = \frac{\partial S_1}{\partial \ell} + \frac{\partial S_2}{\partial \ell} + \ldots$

$$\frac{\partial H_1(L',\ell)}{\partial L} = \lim_{h \to 0} \frac{\partial H_1(L'+h,\ell)}{\partial L} \; . \qquad (29\text{-}29)$$

See the discussion of equation (29-20).

In celestial mechanics one usually starts with Delaunays canonical variables. Given the Hamiltonian $H_o = \frac{\mu^2}{2L^2}$ we transform

$$H(L,G,H,\ell,g,h) \to H^*(L',G',H',-,g,-) \qquad (29\text{-}30)$$

using $S = S_0 + S_1 + S_2 + \ldots$

then transform

$$H^*(L',G',H',-,g,-) \to H^{**}(L'',G'',H'',-,-,-) \qquad (29\text{-}31)$$

using $S^* = S_0{}^* + S_1{}^* + S_2{}^* + \ldots$

Note the double use of the symbol H. For this reason celestial mechanics call the Hamiltonian -F and reserve H for Delauneys' variable.

The short periodic part of the solution involving ℓ and h (equal respectively to n $(t-\tau_o)$ and Ω) is obtained by solving the series of partial differential equations for S. The long periodic terms involving g $(g = \omega)$ come from S*. The secular terms or "mean variables" are obtained from the resulting Hamilton-Jacobi equation.

$$\begin{aligned} \frac{d\ell''}{dt} &= -\frac{\partial H^{**}}{\partial L''} = k_1 & \frac{dL''}{dt} &= \frac{\partial H^{**}}{d\ell''} = 0 \\ \frac{dg''}{dt} &= -\frac{\partial H^{**}}{\partial G''} = k_2 & \frac{dG''}{dt} &= \frac{\partial H^{**}}{\partial g''} = 0 \qquad (29\text{-}32) \\ \frac{dh''}{dt} &= -\frac{\partial H^{**}}{\partial H''} = k_3 & \frac{dH''}{dt} &= \frac{\partial H^{**}}{\partial h''} = 0 \; . \end{aligned}$$

The details for the oblate earth are carried out by Brouwer in his article in Astronomical Journal, Vol 64, pages 378-397 (1959). They are also developed in the book, "Methods of Celestial Mechanics" by D. Brouwer and G. Clemence, Academic Press, 1961. The original work of von Ziepel was for minor planet theory and is found in Arkiv. Mat. Astron. Physik, Vol 11, No. 1, 1916.

In order to illustrate the procedure let us follow Hutcheson ["A Basic Approach to the Use of Canonical Variables," Rand Memo RM-4044-PR, May 1964, AD-600413] and consider a nonlinear spring given by $\ddot{x} + k^2x + \varepsilon x^3 = 0$. For $\varepsilon = 0$ this becomes $\ddot{x} + k^2x = 0$ which represents a unit mass on the end of a spring with spring constant k^2. We solved this kind of problem in Section 13. Consider the problem again.

$$T = \frac{1}{2}\dot{x}^2 \qquad V_o = \frac{k^2}{2}x^2 \qquad L_o = \frac{1}{2}\dot{x}^2 - \frac{k^2}{2}x^2$$

$$y = \frac{\partial L}{\partial \dot{x}} = \dot{x} \tag{29-33}$$

$$H_o = y\dot{x} - \frac{1}{2}\dot{x}^2 + \frac{1}{2}k^2x^2 = \frac{1}{2}y^2 + \frac{1}{2}k^2x^2$$

and hence

$$\frac{dx}{dt} = \frac{\partial H_o}{\partial y} = y \qquad \frac{dy}{dt} = -\frac{\partial H_o}{\partial x} = -k^2x$$

so that

$$\frac{d^2x}{dt^2} = \frac{dy}{dt} = -k^2x \qquad \text{or} \qquad \frac{d^2x}{dt^2} + k^2x = 0 \tag{29-34}$$

which is where we started. The solution is

$$x = c_1 \sin kt + c_2 \cos kt. \tag{29-35}$$

Now let's change to a new coordinate system, a canonical system. We change from (x,y) to (L,ℓ) where L and ℓ are action and angle variables respectively. To do this we use the generating function

$$S = S(x,\ell) = -\frac{1}{2} k\, x^2 \tan \ell \tag{29-36}$$

This is similar to the problem considered in sections 12 and 13.

$$L = -\frac{\partial S}{\partial \ell} = \frac{1}{2} k\, x^2 \frac{1}{\cos^2 \ell} \quad \text{or} \quad x^2 = \frac{2L}{k} \cos^2 \ell \tag{29-37}$$

$$x = \sqrt{\frac{2L}{k}} \cos \ell \tag{29-38}$$

and similarly

$$y = \frac{\partial S}{\partial x} = -kx \tan \ell = -k \frac{\sin \ell}{\cos \ell} \left[\sqrt{\frac{2L}{k}} \cos \ell\right] = -\sqrt{2kL} \sin \ell. \tag{29-39}$$

The transformation is therefore canonical with

$$x = \sqrt{\frac{2L}{k}} \cos \ell \qquad y = -\sqrt{2kL} \sin \ell$$

$$K = H_o + \frac{\partial S}{\partial t} = H_o \tag{29-40}$$

$$\frac{dL}{dt} = -\frac{\partial H_o}{\partial \ell} \qquad \frac{d\ell}{dt} = \frac{\partial H_o}{\partial L}$$

But,

$$H_o = \frac{1}{2} y^2 + \frac{1}{2} k^2 x^2 = \frac{1}{2} [2kL \sin^2 L] + \frac{k^2}{2} [\frac{2L}{k} \cos^2 \ell] = kL \tag{29-41}$$

$$H_o = kL$$

so that

$$\frac{dL}{dt} = -\frac{\partial H_o}{\partial \ell} = 0 \qquad L = \text{constant.}$$

$$\frac{d}{dt} = \frac{\partial H_o}{\partial L} = k \qquad \ell = kt + \text{constant} = k(t-t_o). \tag{29-42}$$

For convenience we select $t_o = 0$, then since

$$x^2 = \frac{2L}{k} \cos^2 \ell \quad \text{we have} \quad x^2(0) = \frac{2L}{k} \quad \text{or} \tag{29-43}$$

$$L = \frac{kx^2(0)}{2} = \text{value of the constant.}$$

To summarize we have for $\ddot{x} + k^2x = 0$

$$H_o(L,\ell) = kL; \qquad L = \frac{1}{2} kx^2(0) - \text{constant};$$

$$\frac{dL}{dt} = -\frac{\partial H_o}{\partial \ell} = 0; \quad \frac{d\ell}{dt} = \frac{\partial H_o}{\partial L} = k; \quad \ell = kt. \tag{29-44}$$

Let's stop a minute and reflect upon this development. How did we pick the particular S function (equation 29-36) to make the canonical transformation? Well we assume the complete solution to the zero order or linear problem is known, e.g., with $\varepsilon = 0$, $\ddot{x} + k^2x = 0$ has the solution

$$x = c_1 \cos kt + c_2 \sin kt$$

$$y = kc_2 \cos kt - kc_1 \sin kt. \tag{29-45}$$

We could then form L and ℓ from a knowledge of the solution since we want ℓ of the form kt and L = constant. For our particular case we have chosen $\ell = k\,(t-t_o)$ and picked $t_o = 0$. This corresponds to the solution,

$$x = c_1 \cos kt$$
$$y = -kc_1 \sin kt. \tag{29-46}$$

First recognizing that $\ell = kt$ we can write these as

$$x - c_1 \cos \ell$$

$$y = -kc_1 \sin \ell$$

then we eliminate c_1 by

$$c_1 = \frac{x}{\cos \ell}$$

and hence

$$y = -k\left(\frac{x}{\cos \ell}\right) \sin \ell = -kx \tan \ell\,.$$

To be canonical we must have

$$y = \frac{\partial S}{\partial x} = -kx \tan \ell$$

which can be integrated to give

$$S = -\frac{k}{2} x^2 \tan \ell.$$

We could also determine S by noting that this particular S generates the action-angle variables and can therefore be computed from

$$S = \int y dx = \text{the action.}$$

See for example Section 62, Chapter 11, page 188, of the second edition of "Classical Mechanics" by Corbin and Stehle.

For our case

$$H = E = \frac{1}{2} y^2 + \frac{1}{2} k^2x^2$$

and hence

$$y = \sqrt{2E - k^2x^2}$$

so that

$$S = \int \sqrt{2E - k^2x^2}\, dx \tag{29-47}$$

where we need to express E in terms of x and ℓ (=kt).

To do this note that

$$E = \frac{1}{2} y^2 + \frac{k^2}{2} x^2 = \frac{k^2x^2}{2} [1 + \frac{y^2}{k^2x^2}]$$

$$E = \frac{k^2x^2}{2} [1 + \frac{k^2c_1^2 \sin^2 kt}{k^2c_1^2 \cos^2 kt}] = \frac{k^2x^2}{2} [1 + \tan^2 kt] = \frac{k^2x^2}{2} \sec^2 kt.$$

Substituting for E in (29-47) gives

$$S = \int kx \sqrt{\sec^2 kt-1}\, dx = \pm k \int x \tan kt\, dx$$

$$S = - \frac{k}{2} x^2 \tan kt \tag{29-49}$$

and then we must recognize that kt is in the form of the angle variable ℓ, i.e., ℓ = kt and hence

$$S = - \frac{k}{2} x^2 \tan \ell = S(x,\ell) \tag{29-50}$$

Since $\ell = kt$, the S function is periodic in time with period $\frac{2\pi}{k}$, i.e., it has the same fundamental frequency as the x variable. This is one of our requirements.

In addition to this approach we could also find L and ℓ by first finding an $\overline{S}$ from the Hamilton-Jacobi partial differential equation $(H + \frac{\partial \overline{S}}{\partial t} = 0)$ for the linear system as in Section 13. We thereby obtain canonical constants α_i and β_i. If we then transform (α_i, β_i) to Delaunay variables, we would have the desired action-angle variables. See Section 17. Perhaps the easiest approach is to find the action-angle variables directly. See Corbin and Stehle referenced above.

Now let us return to the problem. We have a nonlinear spring given by

$$\ddot{x} + k^2 x + \varepsilon x^3 = 0 \, . \tag{29-51}$$

For $\varepsilon = 0$ this reduces to $\ddot{x} + k^2 x = 0$ for which we have

$$H_o = \frac{1}{2} y^2 + \frac{1}{2} k^2 x^2 \, . \tag{29-52}$$

This may be transformed into action-angle variables by using

$$\begin{aligned} x &= \sqrt{\frac{2L}{k}} \cos \ell \\ y &= - \sqrt{2kL} \sin \ell \end{aligned} \tag{29-53}$$

to give

$$H_o(L, \ell) = kL \qquad L = \frac{k\, x^2(0)}{2} \tag{29-54}$$

For the nonlinear system we can easily show that

$$H(x,y) = \frac{1}{2} y^2 + \frac{1}{2} k^2x^2 + \frac{\varepsilon x^4}{4} \tag{29-55}$$

Using the canonical transformation equations, (29-53), this becomes

$$H(L,\ell) = kL + \frac{\varepsilon L^2}{k^2} \cos^4 \ell. \tag{29-56}$$

We now wish to transform H (L,ℓ) into $H^*(L',-)$ where the dash indicates the absence of ℓ'.

When we have done this, we will have

$$\frac{dL'}{dt} = \frac{\partial H^*(L',-)}{\partial \ell'} = 0; \quad \frac{d\ell'}{dt} = \frac{\partial H^*(L',-)}{\partial L'} = \text{constant}. \tag{29-57}$$

From this we see that the time variation in ℓ' is secular and L' will be a constant of the motion. To make the transformation we use the generating function

$$L = \frac{\partial S}{\partial \ell} \qquad \ell' = \frac{\partial S}{\partial L'} \qquad S = S(L',\ell)$$

and expand S as $S = L'\ell + S_1 + S_2 + \ldots.$

First we can write

$$H(L,\ell) = H_o(L) + [H(L,\ell) - H_o(L)] = H_o(L) + H_1(L,\ell)$$

$$H(L,\ell) = kL + \frac{\varepsilon L^2}{k^2} \cos^4 \ell = H_o(L) + H_1(L,\ell)$$

Hence

$$H_1(L,\ell) = H(L,\ell) - H_o(L) = \frac{\varepsilon L^2}{k^2} \cos^4 \ell \tag{29-58}$$

We then proceed to solve equations (29-25) to (29-27). First consider (29-25)

$$H_o(L') = H_o^*(L'). \tag{29-59}$$

Recall

$$H_o\ (L) = kL$$

and hence

$$H_o(L') = kL' = H_o^*(L'). \tag{29-60}$$

The next equation in the von Zeipel set is equation (29-26). Here we need $H_1\ (L',\ell)$ and $\dfrac{\partial H_o(L')}{\partial L}$.

Since $H_o(L) = kL$ then

$$\frac{\partial H_o(L)}{\partial L} = k \tag{29-61}$$

and therefore

$$\frac{\partial H_o(L')}{\partial L} = k \tag{29-62}$$

In addition

$$H_1(L,\ell) = \frac{\varepsilon L^2}{k^2} \cos^4 \ell \tag{29-63}$$

and hence

$$H_1(L',\ell) = \frac{\varepsilon (L')^2}{k^2} \cos^4 \ell \tag{29-64}$$

The von Zeipel equation (29-26) can then be written

$$\left(\frac{\partial S_1}{\partial \ell}\right)\left(\frac{\partial H_o(L')}{\partial L}\right) + H_1(L',\ell) = H_1^*(L')$$

$$\frac{\partial S_1}{\partial \ell}(k) + \frac{\varepsilon(L')^2}{k^2}\cos^4 \ell = H_1^*(L'). \qquad (29\text{-}65)$$

Now recall that $H_1{}^*(L')$ is a constant since our final result will have

$$\frac{dL'}{dt} = -\frac{\partial H^*(L')}{\partial \ell'} = 0 \qquad (29\text{-}66)$$

and therefore L' = constant and hence $H_1{}^*(L')$ = constant. We can use this information to separate and solve equation (29-65).

The $H_1(L',\ell) = \frac{\varepsilon(L')^2}{k^2}\cos^4 \ell$ can be expressed as the sum of two terms, one a constant or non-periodic term in ℓ and the other part periodic in ℓ. For this example we can do this easily since

$$\cos^4 \ell = \frac{3}{8} + \frac{1}{2}\cos 2\ell + \frac{1}{8}\cos 4\ell\ . \qquad (29\text{-}67)$$

Thus we write

$$H_1(L',\ell) = H_{1S} + H_{1P} \qquad (29\text{-}68)$$

$$H_{1S} = \frac{3\varepsilon(L')^2}{8k^2} \qquad H_{1P} = \frac{\varepsilon(L')^2}{2k^2}\left(\cos 2\ell + \frac{1}{4}\cos 4\ell\right) \qquad (29\text{-}69)$$

If we had a more complicated expression for $H_1(L', \ell)$, the secular part could be determined by

$$H_{1S} = \frac{\varepsilon(L')^2}{2\pi k^2} \int_0^{2\pi} \cos^4\ell \, d\ell = \frac{3\varepsilon(L')^2}{8k^2}$$

$$H_{1P} = H_1 - H_{1S}. \tag{29-70}$$

Now returning to equation (29-65) (which came from equation (29-26)) we can always write this in the form,

$$\left(\frac{\partial H_o(L')}{\partial L}\right)\left(\frac{\partial S}{\partial \ell}\right) + H_{1S} + H_{1P} = H_1^* \tag{29-71}$$

and this separates into two equations. This will always occur.

$$H_{1S} = H_1^*(L') = \frac{3\varepsilon(L')^2}{8k^2} \tag{29-72}$$

$$H_{1P} = -\frac{\partial H_o(L')}{\partial L}\frac{\partial S_1}{\partial \ell} = -k\frac{\partial S_1}{\partial \ell}. \tag{29-73}$$

Equation (29-72) gives the first order term in the new Hamiltonian. Using equation (29-69) appropriately for the left hand side of (29-73), equation (29-73) may be written

$$\frac{\partial S_1}{\partial \ell} = -\frac{\varepsilon(L')^2}{2k^3}(\cos 2\ell + \frac{1}{4}\cos 4\ell \tag{29-74}$$

which integrates to give

$$S_1 = -\frac{\varepsilon(L')^2}{4k^3}[\sin 2\ell + \frac{1}{8}\sin 4\ell] + \text{constant} \tag{29-75}$$

Again the constant of integration is of no concern as we deal only with the derivatives of S_1.

From (29-75) we can now begin to construct part of the solution which is periodic in ℓ for the first order, i.e., to first order in ε we can now write

$$L = \frac{\partial S}{\partial \ell} = \frac{\partial(S_o+S_1)}{\partial \ell} = L' - \frac{\varepsilon(L')^2}{2k^3}\left(\cos 2\ell + \frac{1}{4}\cos 4\ell\right) \qquad (29\text{-}76)$$

$$\ell' = \frac{\partial S}{\partial L'} = \frac{\partial(S_o+S_1)}{\partial L'} = \ell - \frac{\varepsilon L'}{2k^3}\left(\sin 2\ell + \frac{1}{8}\sin 4\ell\right). \qquad (29\text{-}77)$$

If first order accuracy is sufficient in calculating the periodic terms, there is no need to determine S_2. At this point the new Hamiltonian H^* is given by (see equations 29-60 and 29-72).

$$\begin{aligned} H^*(L') &= H_o^* + H_1^* + \ldots \\ H^*(L') &= kL' + \frac{3\varepsilon(L')^2}{8k^2} + \ldots \end{aligned} \qquad (29\text{-}78)$$

and since $H^*(L',-)$ is independent of ℓ we have

$$\frac{dL'}{dt} = -\frac{\partial H^*}{\partial \ell} = 0$$

from which we see that L' is a constant. For ℓ' we can write

$$\begin{aligned} \frac{d\ell'}{dt} &= \frac{\partial H^*}{\partial L'} = \frac{\partial H_o^*}{\partial L'} + \frac{\partial H_1^*}{\partial L'} + \ldots \\ \frac{d\ell'}{dt} &= k + \frac{3\varepsilon L'}{4k^2} + \ldots. \end{aligned} \qquad (29\text{-}79)$$

which integrates to

$$\ell' = \left(k + \frac{3\varepsilon L'}{4k^2}\right)(t-t_o). \qquad (29\text{-}80)$$

Recall (29-53),

$$x = \sqrt{\frac{2L}{k}} \cos \ell \qquad y = -\sqrt{2kL} \sin \ell. \tag{29-81}$$

Using (29-76) for L we have

$$x = \sqrt{\frac{2}{k} [L' - \frac{\varepsilon(L')^2}{2k^3} (\cos 2\ell + \frac{1}{4} \cos 4\ell)]} \cos \ell$$
$$\tag{29-82}$$
$$y = -\sqrt{2k [L' - \frac{\varepsilon(L')^2}{2k^3} (\cos 2\ell + \frac{1}{4} \cos 4\ell)]} \sin \ell$$

Solving equation (29-77) gives

$$\ell = \ell' + \frac{\varepsilon L'}{2k^3} (\sin 2\ell + \frac{1}{8} \sin 4\ell) \tag{29-83}$$

and since ℓ is within ε of ℓ' we could approximate this with

$$\ell = \ell' + \frac{\varepsilon L'}{2k^3} (\sin 2\ell' + \frac{1}{8} \sin 4\ell') \tag{29-84}$$

and from (29-80)

$$\ell' = (k + \frac{3\varepsilon L'}{4k^2}) (t-t_o) \tag{29-85}$$

$$L' = \text{constant} .$$

This gives a solution to first order.

The higher order terms are easily found. For example $H_2{}^*(L',-)$ may be obtained from equation (29-27), the next von Zeipel equation.

$$\frac{\partial S_2}{\partial \ell} \frac{\partial H_o(L')}{\partial L} + \frac{1}{2} \left(\frac{\partial S_1}{\partial \ell}\right)^2 \frac{\partial^2 H_o(L')}{\partial L^2} + (\frac{\partial S_1}{\partial \ell}) \frac{\partial H_1(L',\ell)}{\partial L} = H_2{}^*(L'). \tag{29-27}$$

From (29-54) we have $H_o(L) = kL$ and hence $\frac{\partial^2 H_o(L')}{\partial L^2} = 0$. $\frac{\partial S_1}{\partial \ell}$ is given by equation (29-74) as

$$\frac{\partial S_1}{\partial \ell} = -\frac{\varepsilon(L')^2}{2k^3}(\cos 2\ell + \frac{1}{4}\cos 4\ell).$$

From equation (29-64) we have

$$\frac{\partial H_1(L')}{\partial L} = \frac{2\varepsilon L'}{k^2}\cos^4 \ell$$

These reduce equation (29-27) to the form

$$\frac{\partial S_2}{\partial \ell}(k) + [-\frac{\varepsilon(L')^2}{2k^3}(\cos 2\ell + \frac{1}{4}\cos 4\ell)]\,[\frac{2\varepsilon L'}{k^2}\cos^4\ell] = H_2{}^*(L') \quad (29\text{-}86)$$

Since $H_2{}^*(L')$ is a constant and S_2 is strictly periodic in ℓ, we must have

$$H_2{}^*(L') = \text{secular part of } \frac{\partial H_1(L',\ell)}{\partial L}\frac{\partial S_1}{\partial \ell} \qquad (29\text{-}87)$$

which is to say we must determine the non-periodic part of

$$\frac{2\varepsilon^2(L')^3}{k^5}\cos^4 \ell\,[\frac{3}{8} - \cos^4 \ell] \qquad (29\text{-}88)$$

{Recall $\cos^4 \ell = \frac{3}{8} + \frac{1}{8}\cos 4\ell + \frac{1}{2}\cos 2\ell$ and hence

$$\cos 2\ell + \frac{1}{4}\cos 4\ell = 2\,(\cos^4\ell - \frac{3}{8})\}$$

Considering the trigonometric series for $\cos^4 \ell$ and $\cos^8 \ell$ we can write

$$\text{secular part of } \frac{3}{8}\cos^4 \ell = \frac{9}{64}$$

$$\text{secular part of } (-\cos^8 \ell) = -(\frac{9}{64} + \frac{1}{8} + \frac{1}{128}) = -\frac{35}{128}$$

and hence from (29-87) we have

$$H_2^* (L') = - \frac{17}{64} \frac{\varepsilon^2 (L')^3}{k^5} . \tag{29-89}$$

Thus H^* is now given to second order in ε as

$$H^* = kL' + \frac{3\varepsilon (L')^2}{8k^2} - \frac{17}{64} \frac{\varepsilon^2 (L')^3}{k^5} . \tag{29-90}$$

The secular part of the solution equation (29-79) may now be extended to

$$\frac{d\ell'}{dt} = \frac{\partial H^*}{\partial L'} = k + \frac{3\varepsilon L'}{4k^2} - \frac{51}{64} \frac{\varepsilon^2 (L')^2}{k^5} \tag{29-91}$$

$$\frac{dL'}{dt} = - \frac{\partial H^*}{\partial \ell} = 0 . \tag{29-92}$$

From these we now have L' = constant and

$$\ell' = [k + \frac{3\varepsilon L'}{4k^2} - \frac{51}{64} \frac{\varepsilon^2 (L')^2}{k^5}] (t - t_o) . \tag{29-93}$$

From (29-76)

$$L = L' - \frac{\varepsilon (L')^2}{2k^3} [\cos 2\ell + \frac{1}{4} \cos 4\ell] . \tag{29-94}$$

From (29-83)

$$\ell = \ell' + \frac{\varepsilon L'}{2k^3} (\sin 2\ell + \frac{1}{8} \sin 4\ell) . \tag{29-95}$$

Since ℓ and ℓ' are within ε of each other (equation 29-13), we can approximate ℓ by ℓ' in the trigonometric terms to give

$$L = L' - \frac{\varepsilon (L')^2}{2k^3} (\cos 2\ell' + \frac{1}{4} \cos 4\ell') \tag{29-96}$$

$$\ell = \ell' + \frac{\varepsilon L'}{2k^3} (\sin 2\ell' + \frac{1}{8} \sin 4\ell'). \qquad (29\text{-}97)$$

Then using the transformations (29-81) we have

$$x = \sqrt{\frac{2}{k} [L' - \frac{\varepsilon(L')^2}{2k^3} (\cos 2\ell' + \frac{1}{4} \cos 4\ell')} \quad \cos \ell$$

$$y = -\sqrt{2k\ [L' - \frac{\varepsilon(L')^2}{2k^3} (\cos 2\ell' + \frac{1}{4} \cos 4\ell')]} \quad \sin \ell$$

$$\ell = \ell' + \frac{\varepsilon L'}{2k^3} (\sin 2\ell' + \frac{1}{8} \sin 4\ell')$$

$$\ell' = [k + \frac{3\varepsilon L'}{4k^2} - \frac{51}{64} \frac{\varepsilon^2 (L')^2}{k^5}] (t-t_o) .$$

This is the solution, to first order, of the equation $\ddot{x} + k^2 x + \varepsilon x^3 = 0$. In theory the solution may be extended indefinitely by collecting higher order terms in the Taylor's series expansion of H and solving the resulting linear partial differential equations for S_2, S_3, etc. The details are carried out in the referenced RAND report by Hutcheson.

The techniques for solving first order linear partial differential equations are those of Lagranges method and the method of characteristics. See for example, "Partial Differential Equations" by F. H. Miller, Wiley 1941 or "Elements of Partial Differential Equations" by I. N. Sneddon, McGraw-Hill, 1957, page 44, et. seq.

To clearly understand the technique the student should solve the equation

$$\ddot{x} + k^2 x + \varepsilon x^5 = 0$$

for $x = x(t)$ to order ε^3 in the periodic terms and through ε^4 in the secular part. We will compare this result with numerical integration.

(Hint: show that $\quad H(L,\ell) = kL + \frac{\varepsilon}{6}[\frac{8L^3}{k^3}\cos^6\ell]$)

Recall we replaced ℓ by ℓ' in the trigonometric functions in equations (29-94) and (29-95). If we desire to be more precise we can invert the series expansion, such as (29-95), in order to solve for ℓ as a series expansion in ℓ'. Given

$$\ell = \ell' + \frac{\varepsilon L'}{2k^3}(\sin 2\ell + \frac{1}{8}\sin 4\ell)$$

we can find $\ell = f(\ell', \varepsilon, L')$. This is done by using Lagrange's formula. See Brown and Shook, "Planetary Theory", Dover, page 37, et. seq.

In general if we are given

$$y = x + \varepsilon\phi(y)$$

where ε is a small parameter and ϕ and its derivatives are continuous functions of y, and we want the expansion of some other function of y such as $F = F(y)$ in powers of ε with coefficients which are functions of x alone, we may use Lagrange's theorem. This theorem states that the answer to our problem is given by the series,

$$F(y) = F_x + \varepsilon(\phi_x \frac{dF_x}{dx}) + \frac{\varepsilon^2}{2!}\frac{d}{dx}(\phi_x^2 \frac{dF_x}{dx}) + \frac{\varepsilon^3}{3!}\frac{d^2}{dx^2}(\phi_x^3 \frac{dF_x}{dx}) + \ldots$$

where $F_x = F(x)$, $\phi_x = \phi(x)$, i.e., replace y by x. In particular when $F(y) = y$ we have

$$y = x + \varepsilon\phi_x + \frac{\varepsilon^2}{2!}\frac{d}{dx}(\phi_x^2) + \frac{\varepsilon^3}{3!}\frac{d^2}{dx^2}(\phi_x^3) + \ldots$$

This theorem may be extended to several variables. We can then invert any series expansion into some other series. This is a very handy and oft used tool in Celestial Mechanics. See Brown and Shook[5] or Smart[6] for further details.

As an example of the procedure consider Kepler's equation

$$E - e \sin E = M$$

This may be written as

$$E = M + e \sin E = M + e\phi(E)$$

which is in the required form with $y = E$ and $x = M$. Applying Lagrange's formula gives the following:

$$E = M + e\phi_M + \frac{e^2}{2!}\frac{d}{dM}[\phi_M^2] + \frac{e^2}{3!}\frac{d^2}{dM^2}[\phi_M^3] + \ldots$$

ϕ_M = sinE with $E \to M$, i.e.,

$\phi_M = \sin M$

Hence

$$E = M + e \sin M + \frac{e^2}{2!} [2 \sin M \cos M] + \frac{e^3}{3!} [3 \sin^2 M \cos M] + \ldots$$

$$E = M + e \sin M + \frac{e^2}{2!} \sin 2M + \frac{e^3}{3!2^2} (3^2 \sin 3M - 3 \sin M) + \ldots$$

which corresponds to equation (4-15) on page 33 and reduces to

$$E = M + (e - \frac{e^3}{8} + \ldots) \sin M$$

$$+ (\frac{e^2}{2} + \ldots) \sin 2M + (\frac{3e^3}{8} \ldots) \sin 3M + \ldots \text{ etc.}$$

The von Zeipel method is closely related to that developed by Lindstedt and by Delaunay except that these methods do not make use of a generating function. See Chapter IX of An Introductory Treatise on the Lunar Theory by E. W. Brown. The use of a generating function greatly simplifies the procedure.

Giacaglia[1] points out that the concept of adiabatic invariants in Quantum Mechanics is quite analogous to the concept of "mean variables" in von Zeipel's method, or to a certain extent to what Whittaker calls Adelphic Integrals.

Morrison in AIAA Preprint No. 65-687 (September 1965) has compared the generalized method of averages with that of von Zeipel. He shows the two methods lead to the same results; however, the von Zeipel method is most convenient when the canonical variables are used.

The generalized method of averages is a generalization of the method of Section 23 to include higher order terms. A very good discussion of the method of averages, how it gives rise to an apparent

secular term and how that can be eliminated, can be found in Chapter one of the book by Chihiro Hayaski[8] entitled, "Nonlinear Oscillations in Physical Systems." His section 1.2 starting on page 13 develops the same Duffing's equation as we have covered here.

The method of averages assumes a solution $x(\tau)$ in a power series with respect to some small parameter ϵ with coefficients in the series being periodic functions of $\tau = \omega t$. Thus

$$x(\tau) = x_0(\tau) + \epsilon x_1(\tau) + \epsilon^2 x_2(\tau) + \dots \quad , \tag{29-98}$$

the $x_i(\tau)$ being functions of τ of period 2π. In order to eliminate secular terms we also develop a second unknown quantity ω with respect to ϵ, e.g.,

$$\omega = \omega_0 + \epsilon\omega_1 + \epsilon^2 \omega_2 + \dots \tag{29-99}$$

We then substitute (29-98) and (29-99) into the differential equation and equate the coefficients of like powers of ϵ. This gives a sequence of differential equations in $x_i(\tau)$ which also involve the ω_i quantities. Since only periodic solutions are under consideration (a conservative system) we choose the $\tau = 0$ point arbitrarily to give initial conditions of $\dot{x}(\tau) = 0$ at $\tau = 0$. For example consider

$$\frac{d^2x}{dt^2} + x + \epsilon x^3 = 0$$

Change the independent variable t to $\tau = \omega t$.

$$\omega^2 \frac{d^2x}{d\tau^2} + x + \epsilon x^3 = \omega^2 \ddot{x} + x + x^3 = 0 \tag{29-100}$$

Substituting (29-98) and (29-99) into (29-100) and equating the coefficients of like powers of ε leads to the following sequences of linear equations:

$$\varepsilon^0: \qquad \omega_0^2 \ddot{x}_0 + x_0 = 0 \tag{29-101}$$

$$\varepsilon^1: \qquad \omega_0^2 \ddot{x}_1 + x_1 = -2\omega_0 \omega_1 \ddot{x}_0 - x_0^3 \tag{29-102}$$

$$\varepsilon^2: \qquad \omega_0^2 \ddot{x}_2 + x_2 = -(2\omega_0\omega_2 + \omega_1^2)\ddot{x}_0 - 2\omega_0\omega_1 \ddot{x}_1 - 3x_0^2 x_1 \tag{29-103}$$

The initial conditions are assumed to be

$$x(0) = A; \frac{dx(0)}{dt} = 0 \quad ; \quad \frac{dx(0)}{d\tau} = 0 \ .$$

and since $x(\tau + 2\pi) = x(\tau)$ we have the following conditions:

$$x_i(\tau + 2\pi) = x_i(\tau) \tag{29-104}$$

$$\begin{aligned} & x_0(0) = A \qquad x_{i+1}(0) = 0 \\ & \dot{x}_i(0) = 0 \qquad i = 0, 1, 2, \ldots \end{aligned} \tag{29-105}$$

Solving (29-101) with the use of these conditions gives

$$x_0 = A \cos \tau \tag{29-106}$$

$$\omega_0 = 1 \tag{29-107}$$

Then using (29-106) and (29-107), equation (29-102) becomes

$$\ddot{x}_1 + x_1 = (2\omega_1 - \frac{3}{4} A^2) A \cos \tau - \frac{1}{4} A^3 \cos 3\tau. \qquad (29\text{-}108)$$

If the coefficient of $\cos \tau$ were not zero, the solution of (29-108) would contain a term of the form, $\tau \sin \tau$, i.e., a secular term. The periodicity condition for $x_1(\tau)$ therefore requires that the coefficient of $\cos \tau$ be zero, i.e.,

$$\omega_1 = \frac{3}{8} A^2.$$

Hence using (29-105) the solution of (29-108) becomes

$$x_1 = \frac{1}{32} A^3(- \cos \tau + \cos 3\tau).$$

By proceeding analogously, we obtain

$$\omega_2 = -\frac{21}{256} A^4$$

$$x_2 = \frac{23}{1024} A^5 \cos \tau - \frac{3}{128} A^5 \cos 3\tau + \frac{1}{1024} A^5 \cos 5\tau .$$

The solution up to terms of second order in ε becomes

$$x(t) = (A - \frac{1}{32}\varepsilon A^3 + \frac{23}{1024} \varepsilon^2 A^5] \cos\omega t$$

$$+ (\frac{1}{32} \varepsilon A^3 - \frac{3}{128} \varepsilon^2 A^5) \cos 3\omega t + \frac{1}{1024} \varepsilon^2 A^5 \cos 5\omega t + \ldots$$

and

$$\omega = 1 + \frac{3}{8} \varepsilon A^2 - \frac{21}{256} \varepsilon^2 A^4 + \ldots$$

as before, the frequency ω depends on the amplitude A.

REFERENCES

1. Giacoglia, G., "Notes on von Zeipel's Method", NASA X-547-64-16, 1964.

2. Kyner, W. T., Mathematical Theory of Orbit about an Oblate Planet, SIAM Journal, Vol 13, No. 1, p 136, 1965.

3. Struble, R. A., "The Geometry of the Orbits of Artificial Satellites", Arch. Rational Mech. Anal., Vol 7, p 87, 1961.

4. Cole, J. D. and J. Kevorkian, "Uniformly Valid Approximations for Certain Nonlinear Differential Equations," Proceeds of Symposium on Nonlinear Differential Equations, Academic Press, 1963, pp 113-120.

5. Brown, E. W. and C. A. Shook, "Planetary Theory," Dover, 1933 (Reprint 1964), Chapter II, p 36.

6. Smart, W. M., "Celestial Mechanics," Longmans, 1953, Chapter 3, p 26.

7. J. H. Hutcheson, "A Basic Approach to the Use of Canonical Variables," RAND Memo RM-4074-PR, May 1964, AD-600413.

8. Hayashi, Chihiro,"Nonlinear Oscillations in Physical Systems" McGraw-Hill, 1964.

30. SPECIAL PERTURBATIONS

Given the equations of motion in the form

$$\frac{d^2x_i}{dt^2} + \frac{k^2Mx_i}{r_i^3} = \frac{\partial R}{\partial x_i} \tag{30-1}$$

We could apply standard numerical integration techniques and obtain numerical solutions for x_i. This method is called Cowell's method in Celestial Mechanics. It is simple to apply, but has important deficiencies. For one, the elements x_i and $\dot{x}_i$ do not give a very vivid description of the motion. Certainly nothing like listing how the elements of an ellipse are slowly varying. The second problem is that these x_i variables undergo large excursions requiring small integration intervals. This means more integration steps are required for a given time of integration. This greatly increases the round off error. So although Cowell's method is simple to set up, it can be dangerous to use.

Since the orbital elements (elliptical elements) change only slowly, we can numerically integrate the equations of motion given by equation (18-16) on page 176. Here we will be able to take large integration intervals. This scheme is referred to as the variation of parameter technique. It is particularly useful for small but continuously changing perturbations such as the use of micro-thrust rockets, etc.

A third method in the area of special perturbations is Encke's method. Lets illustrate the principle by a simple example.

Consider the case of a circular orbit with the velocity perpendicular to the radius vector. Assume the perturbing force is that due to drag which acts opposite to the velocity. We then write the radial and tangential components of force as

$$\ddot{r} - r\dot{\theta}^2 = -\frac{k^2 M}{r^2} \qquad (30\text{-}2)$$

$$r\ddot{\theta} + 2\dot{r}\dot{\theta} = -\frac{D}{m} \qquad (30\text{-}3)$$

$$D = \text{Drag Force} = \frac{1}{2}\rho\,\frac{C_D A}{m}\,V^2 = \frac{1}{2}\rho\left(\frac{C_D A}{m}\right)(r\dot{\theta})^2 \qquad (30\text{-}4)$$

Now let

$$r = r_o + \Delta r \qquad (30\text{-}5)$$

$$\theta = \theta_o + \Delta\theta \qquad (30\text{-}6)$$

such that r_o and θ_o are defined as solutions to

$$\ddot{r}_o - r_o\dot{\theta}^2 = -\frac{kM}{r_o^2} \tag{30-7}$$

$$r_o\ddot{\theta}_o + 2\dot{r}_o\dot{\theta}_o = 0 \tag{30-8}$$

Thus r_o, θ_o are known functions of time. They are the solution to the unperturbed two body problem. Now substituting (30-5) and (30-6) into (30-2) gives

$$\ddot{r}_o + \Delta\ddot{r} - (r_o + \Delta r)\ (\dot{\theta}_o + \Delta\dot{\theta})^2 = -\frac{k^2M}{(r_o + \Delta r)^2} = -\frac{k^2M}{r_o^2\left(1 + \frac{\Delta r}{r_o}\right)^2}$$

We can expand $\frac{1}{\left(1 + \frac{\Delta r}{r_o}\right)^2}$ in a series expansion and if we neglect terms of order Δr compared to r_o; $\Delta\dot{\theta}$ compared to $\dot{\theta}_o$, then terms like $(\Delta\dot{\theta})^2$, $\Delta r\ \Delta\theta$, are very small indeed. If we neglect these second order terms we have

$$\ddot{r}_o + \Delta\ddot{r} - r_o\dot{\theta}_o^2 - 2r_o\dot{\theta}_o\Delta\dot{\theta} - \dot{\theta}_o^2\Delta r = -\frac{k^2M}{r_o^2}\quad 1 - 2\frac{\Delta r}{r_o} + \dots \tag{30-10}$$

Making use of equation (30-7) this reduces to

$$\ddot{\Delta r} - 2r_o\dot{\theta}_o\Delta\dot{\theta} - \Delta r\left[\frac{2k^2M}{r_o^3} + \dot{\theta}_o^2\right] = 0 \tag{30-11}$$

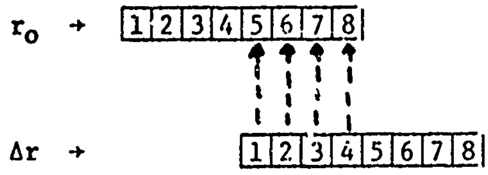
likewise (30-3) reduces to

$$\Delta\ddot{\theta} + \frac{2\dot{\theta}_o}{r_o}\Delta\dot{r} = -\frac{D}{mr_o} = -\frac{1}{2}\rho\dot{\theta}_o^2 r_o\left(\frac{C_DA}{m}\right). \tag{30-12}$$

Equations (30-7) and (30-8) give r_o and θ_o as a function of time and hence (30-11) and (30-12) are linear equations with time varying coefficients. We are solving for Δr and $\Delta\theta$ which will be small and whose variations are presumably slow. Thus large integration intervals may be used. But there is another advantage. The total variable such as r, is composed of $r = r_o + \Delta r$. The digital computer has a finite number of digits in its register, i.e., the read-out has a certain number of significant figures.

|1|2|3|4|5|6|7|8|

Of the total r, the Δr represents only a small portion. Thus Δr which is added to r_o to give r is done as below:



Any round-off in Δr is lost in the formation of r. Encke's method thus gives large integration intervals because Δr ($\Delta\theta$) only is being integrated and it has a bonus by being less affected by round-off errors.

Now let's apply Encke's method to the n-body problem. Consider the x component only. The y and z components are similarly found.

$$\ddot{x} = -\frac{k^2 M_o}{r^3}\, x + k^2 \sum_{i=1}^{n} m_i \left[\frac{x_i - x}{r_i^3} - \frac{x_i}{r_{io}^3}\right] \qquad (30\text{-}13)$$

$$r_i^2 = (x - x_i)^2 + (y - y_i)^2 + (Z - Z_i)^2$$

$$(30\text{-}14)$$

$$r_{io}^2 = x_i^2 + y_i^2 + Z_i^2 \qquad r^2 = x^2 + y^2 + Z^2 \ .$$

Now let $x = x_o + \xi \quad y = y_o + \eta \quad Z = Z_o + \zeta$ with

$$\ddot{x}_o = -\frac{kM_o x_o}{r_o^3} \quad \ddot{y}_o = -\frac{kM_o y_o}{r_o^3} \quad \ddot{Z}_o = -\frac{kM_o Z_o}{r_o^3} \qquad (30\text{-}15)$$

then

$$\ddot{x} - \ddot{x}_o = \xi = k^2 M_o \left(\frac{x_o}{r_o^3} - \frac{x}{r^3}\right) + k^2 \sum_i^n m_i \left(\frac{x_i - x}{r_i^3} - \frac{x_i}{r_{io}^3}\right) \qquad (30\text{-}16)$$

But -

$$\frac{x_o}{r_o^3} - \frac{x}{r^3} = \frac{1}{r_o^3}\left(x_o - \frac{r_o^3}{r^3}x\right) = \frac{1}{r_o^3}\left(x - \xi - \frac{r_o^3}{r^3}x\right)$$

$$= \frac{1}{r_o^3}\left[\left(1 - \frac{r_o^3}{r^3}\right)x - \xi\right] \qquad (30\text{-}17)$$

Recall,

$$r^2 = x^2 + y^2 + Z^2 = (x_o + \xi)^2 + (y_o + \eta)^2 + (Z_o + \zeta)^2$$

$$r^2 = r_o^2 + 2x_o\xi + 2y_o\eta + 2Z_o\zeta + \xi^2 + \eta^2 + \zeta^2$$

$$\frac{r^2}{r_o^{\,2}} = 1 + 2\,\frac{(x_o + \frac{1}{2}\xi)\xi + (y_o + \frac{1}{2}\eta)\eta + (Z_o + \frac{1}{2}\zeta)\zeta}{r_o^{\,2}} \qquad (30\text{-}18)$$

Let

$$q = \frac{(x_o + \frac{1}{2}\xi)\xi + (y_o + \frac{1}{2}\eta)\eta + (Z_o + \frac{1}{2}\zeta)\zeta}{r_o^{\,2}}$$

so that

$$\frac{r^2}{r_o^{\,2}} = 1 + 2q \qquad \frac{r_o^{\,3}}{r^3} = (1 + 2q)^{-3/2} \qquad (30\text{-}19)$$

and the term in equation (30-17) as modified by (30-19) can be written

$$1 - \frac{r_o^{\,3}}{r^3} = 1 - (1 + 2q)^{-3/2} \qquad (30\text{-}20)$$

Of ξ, η, $\zeta \ll x_o$, y_o, Z_o respectively we can neglect their squares giving

$$q \cong \frac{x_o \xi + y_o \eta + Z_o \zeta}{r_o^3} \tag{30-21}$$

but with the use of high speed computers such an approximation is not usually used. In any event, in equation (30-20) we could use the expansion

$$1 - \left(\frac{r_o}{r}\right)^3 = fq = 1 - (1 + 2q)^{-3/2} \cong 3q - \frac{15}{2} q^2 + \ldots \tag{30-22}$$

however, it is more convenient to introduce another variable f defined by

$$f = \frac{1 - (1 + 2q)^{-3/2}}{q} \tag{30-23}$$

For small q, $f \to 3.0$ and f changes much less rapidly than q making a table of f much easier to interpolate. Returning to equation (30-16) we can write this as

$$\ddot{\xi} = \frac{k^2 M_o}{r_o^3} \left[\left(1 - \frac{r_o^3}{r^3}\right) x - \xi\right] + k^2 \sum_i m_i \left(\frac{x_i - x}{r_i^3} - \frac{x_i}{r_{io}^3}\right) \tag{30-24}$$

$$\ddot{\xi} = \frac{k^2 M_o}{r_o^3}\left[fqx - \xi\right] + k^2 \sum_i m_i \left(\frac{x_i - x}{r_i^3} - \frac{x_i}{r_{io}^3}\right) \quad . \tag{30-25}$$

Similar equations are obtained for η and ζ. Thus we are integrating only the perturbations from some nominal elliptical orbit. Six constants of integration are required x_o, y_o, Z_o, $\dot{x}_o$, $\dot{y}_o$, $\dot{Z}_o$ at the epoch of osculation. This nominal orbit (subscripted zero) must be both tangent to and have the same velocity as the x, y, z orbit at time of epoch. Such an orbit is called the osculating orbit. When the orbit is only tangent, it is called an intermediary orbit.

To speed the computations we let

$$fq = 3q - \frac{3\cdot5}{2!}q^2 + \frac{3\cdot5\cdot7}{3!}q^3 \cdots \tag{30-26}$$

with

$$q = \frac{1}{r_o^2}\left[\xi(x_o + \tfrac{1}{2}\xi) + \eta(y_o + \tfrac{1}{2}\eta) + \zeta(Z_o + \tfrac{1}{2}\eta)\right] \tag{30-27}$$

When the q value begins to become too large we shift to a new osculating orbit, i.e., a new set of x_o, y_o, Z_o, $\dot{x}_o$, $\dot{y}_o$, $\dot{Z}_o$ nominal

elliptical orbit. This process is called rectifying the orbit. How do we decide when to rectify? The series

$$1 - \left(\frac{r_o}{r}\right)^3 = fq = 3q - \frac{3\cdot 5}{2!}\, q^2 + \frac{3\cdot 5\cdot 7}{3!}\; q^3 - \cdots = 1 - \frac{1}{(1 + 2q)^{3/2}} \tag{30-28}$$

converges for $-\frac{1}{2} < q < \frac{1}{2}$ and can be represented in general as

$$fq = \sum_{i=1}^{\infty} (-1)^{i-1}\, q^i \,\frac{(2_i + 1)!}{(i!)^2\; 2^i} \tag{30-29}$$

The allowable error in $\ddot{\xi}$ can be written

$$\text{Error }(\ddot{\xi}) = \text{Error}\left[\frac{\mu}{r_o{}^3}\,(fqx - \xi) + k^2 \sum_i m_i \left(\frac{x_i - x}{r_i{}^3} - \frac{x_i}{r_{io}^3}\right)\right] \tag{30-30}$$

$$\doteq \text{Error in}\left(\frac{fq}{r_o{}^2}\right)$$

Therefore

$$r_o{}^2\ (\text{ERROR IN } \ddot{\xi}) \doteq a_{n+1}\, q^{n+1}$$

where a_{n+1} is the coefficient of the first neglected term in the series expansion for f_{q} (see equation 30-29). The limits on q needed for rectification are established as

$$|q| < \sqrt[n+1]{r_0^2 \frac{(\text{error in } \ddot{\xi})}{|a_{n+1}|}} \tag{30-31}$$

$$|a_{n+1}| = \frac{|2n + 2|!}{[(n + 1)!]^2 \, 2^{n+1}} \tag{30-32}$$

Thus we decide how much error to allow in $\ddot{\xi}$, $\ddot{\eta}$ and $\ddot{\zeta}$ and then keep track of $|q|$. When $|q|$ exceeds the limit set by (30-31) we stop the integration and compute x, y, z, $\dot{x}$, $\dot{y}$, $\dot{z}$ and from these values re-establish a new set of values, a new nominal ellipse, to compute a new set x_0', y_0', z_0', $\dot{x}_0'$, $\dot{y}_0'$, $\dot{z}_0'$ and then proceed as before. If we have to rectify very often then Encke's scheme takes a lot of computer time and introduces round-off error as we are forced to convert 8 significant figures for the rectified initial conditions. Thus we throw away the advantages of Encke's method.

Lunar trajectories which are basically two nominal ellipses are easily handled by Encke's method. Microthrust trajectories are continually changing ones and hence best suited to variation in parameter techniques as are the study of effects of solar radiation. High impulse thrust are drastic changes and must use Cowell's method.

31. EXAMPLE OF LINEARIZED ENCKE METHOD

Blitzer and others have used essentially Encke's method to obtain first order results. Consider the effect of ellipticity of the earth's equator on a 24 hour nearly circular satellite. The external potential of the earth can be represented by

$$U = + \frac{\mu}{r}\left[1 - \sum_{n=2}^{\infty} J_n \left(\frac{a_e}{r}\right)^n P_n(\cos\theta) + \sum_{n=2}^{\infty}\sum_{m=1}^{n} J_{n,m}\left(\frac{a_e}{r}\right)^n P_{n,m}(\cos\theta)\cos m(\lambda - \lambda_{n,m})\right]$$

where r is the geocentric distance λ is the geographic longitude, θ is the polar angle, a_e the mean radius in the equatorial plane, $P_{n,m}$ are related Legendre polynomials. The leading terms are

$$U = + \frac{\mu}{r}\left[1 - \frac{J_2 a_e^2}{2r^2}\left(3\cos^2\theta - 1\right) + \frac{J_{22} a_e^2}{2r^2}\sin^2\theta\cos 2(\lambda - \lambda_{22})\right]$$

$\lambda_{n,m}$ is the longitude of the major axis of the equatorial ellipse, thus $\lambda - \lambda_{n,m}$ is longitude measured from the major axis. J_2 and J_{22} are numerical constants of the order of 10^{-3} and 10^{-5} respectively.

Assume the satellite is moving in the same direction as the earth's rotation and in or near the equatorial plane. It is convenient to study the motion in a rotating spherical-polar coordinate system with its origin at the earth's center and rotating with the angular velocity ω of the earth. The kinetic energy can be written as:

$$T = \frac{1}{2}\left[\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,(\dot{\lambda}+\omega)^2\right] \qquad (31\text{-}2)$$

The equations of motion become

$$\ddot{r} - r\dot{\theta}^2 - r\sin^2\theta(\dot{\lambda}+\omega)^2 = -\frac{\mu}{r^2} + \frac{3\mu J_{2} a_e^2}{2r^4}\left(3\cos^2\theta - 1\right) - \frac{3\mu J_{22} a_e^2}{2r^4}\sin^2\theta\cos 2(\lambda-\lambda_{22}) \qquad (31\text{-}3)$$

$$\frac{d}{dt}\left[r^2\sin^2\theta(\dot{\lambda}+\omega)\right] = -\frac{\mu J_{22} a_e^2}{r^3}\sin^2\theta\,\sin 2(\lambda-\lambda_{22}) \qquad (31\text{-}4)$$

$$\frac{d}{dt}(r^2\dot{\theta}) - \frac{r^2(\dot{\lambda}+\omega)^2}{2}\sin 2\theta = \frac{3\mu J_2 a_e^2}{2r^3}\sin 2\theta + \frac{\mu J_{22} a_e^2}{2r^3}\sin 2\theta\cos 2(\lambda-\lambda_{22}) \qquad (31\text{-}5)$$

We are looking for possible stationary states in this rotating frame in order to place a 24 hour satellite. From symmetry, the stationary positions must be on the equator, thus $r = r_0$ = constant, $\theta = \frac{\pi}{2}$, $\lambda = \lambda_0$ = const. For this case equations (31-3) to (31-5) becomes, $(\dot{\theta} = \dot{\lambda} = 0)$

$$r_0\omega^2 = \frac{\mu}{r_0^2} + \frac{3\mu J_2 a_e^2}{2r_0^4} + \frac{3\mu J_{22} a_e^2}{2r_0^4}\cos 2(\lambda_0 - \lambda_{22}) \qquad (31\text{-}6)$$

$$0 = \frac{\mu J_{22} a_e^2}{r_0^3}\sin 2(\lambda_0 - \lambda_{22}) \qquad (31\text{-}7)$$

$$0 = 0 \qquad (31\text{-}8)$$

From (7) we find stationary conditions at $\lambda_0 - \lambda_{22} = 0$ or $\lambda_0 = \lambda_{22}$, $\lambda_0 = \lambda_{22} + \frac{\pi}{2}$, $\lambda_0 = \lambda_{22} + \pi$ and $\lambda_0 = \lambda_{22} + \frac{3\pi}{2}$. For each of these four points there is a unique value of r_0 fixed by (31-6). The four points lie on the major or minor axis of the ellipse. Now let's investigate motion about these four points. To do so we use dimensionless variables.

$$Z = \frac{r}{a_e} \qquad \tau = \omega t \qquad \alpha = \frac{\mu}{\omega^2 a_e{}^3}$$

We then obtain in lieu of (31-3), (31-4) and (31-5)

$$Z'' - Z\theta'^2 - (1+\lambda')^2 Z \sin^2\theta = -\frac{\alpha}{Z^2} + \frac{3\alpha J_2}{2Z^4}\left(3\cos^2\theta - 1\right) - \frac{3\alpha J_{22}}{2Z^4}\sin^2\theta\cos 2(\lambda-\lambda_{22}) \tag{31-9}$$

$$\frac{d}{d\tau}\left[(1+\lambda')Z^2\sin^2\theta\right] = -\frac{\alpha J_{22}}{Z^3}\sin^2\theta\sin 2(\lambda-\lambda_{22}) \tag{31-10}$$

$$\frac{d}{d\tau}\left[Z^2\theta'\right] - \frac{Z^2(1+\lambda')^2}{2}\sin 2\theta = \frac{3\alpha J_2}{2Z^3}\sin 2\theta + \frac{\alpha J_{22}}{2Z^3}\sin 2\theta\cos 2(\lambda-\lambda_{22}). \tag{31-11}$$

where primes denote derivatives with respect to τ. To examine stability we look at motion in the neighborhood of the stationary points determined by (31-6) and (31-7) For small displacements we set

$$\begin{aligned} Z &= Z_0 + \Delta \\ \lambda &= \lambda_0 - \lambda_{22} + \phi \\ \theta &= \frac{\pi}{2} - \delta \end{aligned} \tag{31-12}$$

where $\Delta \ll Z_o \quad \phi \ll \pi \quad \delta \ll \pi$. If we substitute into (31-9, -10, -11) linearize by ignoring square and product terms of Δ, ϕ, δ and their derivatives, we find

$$\Delta'' - (1 + a)\Delta - 2Z_o\phi' = 0 \tag{31-13}$$

$$Z_o^2\phi'' + b\phi + 2Z_o\Delta' = 0 \tag{31-14}$$

$$\delta'' + c^2\delta = 0 \tag{31-15}$$

where

$$a = \frac{2\alpha}{Z_o{}^3} + \frac{6\alpha J_2}{Z_o{}^5} \pm \frac{6\alpha J_{22}}{Z_o{}^5}$$

$$b = \pm \frac{2\alpha J_{22}}{Z_o{}^3} \tag{31-16}$$

$$c^2 = 1 + \frac{3\alpha J_2}{Z_o{}^5} \pm \frac{\alpha J_{22}}{Z_o{}^5}$$

where the upper sign pertains to motion about a stationary point on the major axis and the lower sign for motion about a point on the minor axis.

Note the δ equation (31-15) is decoupled from the others, hence latitude motion is simply periodic with frequency c. Since J_2 and J_{22} are both small and $Z_o \doteq 6.63$, this frequency, c, is practically unity, or in dimensional time, very nearly one sidereal day. The amplitude is equal to the angle of inclination of the orbit to the equator.

For equations (31-13) and (31-14) we assume a periodic solution and write

$$\Delta = A e^{ip\tau} \qquad \phi = B e^{ip\tau} \quad .$$

If these are substituted appropriately into (31-13) and (31-14) we obtain

$$(1 + a + p^2)A + 2ipZ_oB = 0$$

$$2ipZ_oA + (b - p^2Z_o^2)B = 0$$

For this pair to have a nontrivial solution for A and B, the determinant of the coefficients must vanish. This determines the frequency, p, namely,

$$p^2 = \frac{(3Z_o^2 - aZ_o^2 + b) \pm \left[(3Z_o^2 - aZ_o^2 + b)^2 + 4(1+a)bZ_o^2\right]^{1/2}}{2Z_o^2}$$

Since $J_{22} \ll 1$, thence $b \ll 1$, so square of the two frequencies are

$$p_1^2 \cong (3 - a) + \frac{4b}{(3 - a)Z_o^2}$$

$$p_2^2 = -\frac{(1 + a)b}{(3 - a)Z_o^2}$$ to evaluate consider -

$$a \cong \frac{2\alpha}{Z_o^3} = \frac{2\mu}{\omega^2 r_o^3}$$ $c = 1$

$$b \cong \pm \frac{2\alpha J_{22}}{Z_o^3} = \pm \frac{2\mu J_{22}}{\omega^2 r_o^3}$$ and by Kepler's law

$$\omega^2 r_o^3 = \mu$$

hence

$a = 2.0$ $\quad b = \pm 2J_{22}$ $\quad c = 1$ $\quad Z_o \cong 6.63$

and hence

$$p_1^2 = (3 - 2) \pm \frac{4 \cdot 2J_{22}}{(3 - 2)(6.63)^2} = 1 \pm \frac{8J_{22}}{(6.63)^2}$$

$$p_1 \approx 1$$

$$p_2^2 = - \pm \frac{(1 + 2)2J_{22}}{(3 - 2)Z_o^2} = \mp \frac{6J_{22}}{Z_o^2} \quad .$$

Note that in the neighborhood of a stationary point on the major axis (when the upper sign applies), p_2^2 is negative and hence the motion is unstable. On the other hand, near a stationary point on the minor axis

$$p_2^2 = \frac{6J_{22}}{Z_o^2} = 0.137J_{22}$$

so both p_1 and p_2 are real and the motion is oscillatory and hence stable. In actual time the periods are

$$P_1 = \frac{1}{p_1} \text{ days} \approx 1 \text{ day}$$

$$P_2 = \frac{1}{p_2} \text{ days} \approx \frac{2.71}{\sqrt{J_{22}}} \approx 2\ 1/3 \text{ years}$$

Blitzer goes on to study motion near these stationary points and uses a computer to study motion at larger distances. See "Effect of Ellipticity of the Equator on 24 Hour Nearly Circular Satellite Orbits" by Leon Blitzer, E. M. Boughton, G. Kang and R. M. Page, Journal of Geophysical Research, Vol. 67, No. 1, January 1962, p. 329. Frick and Garber obtained analytical solutions for large amplitudes about the stable points. See "Perturbations of a Synchronous Satellite," Rand Corporation R-399, May 1962. They show that a satellite placed at major axis point will drift away at an average rate of 0.45°/day for first 90 degrees. Maximum $\Delta r \approx 27.5$ nautical miles. If injected at any other longitude λ_o, the satellite will have periodic oscillations of amplitude λ_o about the minor axis. For example with $\lambda - \lambda_{22} = 45°$, it drifts to 90° and return with period of 1.78 years and $\Delta r = 19.5$ n.m.

Later when satellite data indicated many higher order J_{nm} coefficients were as large as J_{22}, Blitzer recomputed the problem taking into account all the other J_{nm} terms. See "Equilibrium Positions and Stability of 24 Hour Satellite Orbits," Journal of Geophysical Research, Vol. 70, No. 16, August 1965, p. 3987. He finds there are still only four equilibrium points although they now do not lie exactly on the major-minor axis and the symmetry is destroyed.

As the earth rotates under the satellite the potential field is time varying. When it varies at nearly the same rate as one of the elements of the satellite we have the possibility of resonance. Thus for certain

satellites, depending on their period, certain J_{nm} terms, though small, cause a measurable effect. Recently $J_{15,13}$, $J_{13,13}$, $J_{15,14}$ have been determined this way. See "Observation of Resonance Effects on Satellite Orbits Arising from the 13th and 14th Order Tesseral Gravitational Coefficients," R. J. Anderle, Journal of Geophysical Research, Vol. 70, No. 10, May 1965, p. 2453. A thorough study of this resonance effect can be found in the following references:

S. M. Yionoulis, "Resonant Geodesy," Applied Physics Lab Tech Report TG-633, December 1964. Applied Physics Lab, The Johns Hopkins University, Silver Springs, Maryland.

S. M. Yionoulis, "A Study of the Resonance Effects Due to the Earth's Potential Function," Journal of Geophysical Research, Vol. 70, No. 24, December 1965, p. 5991.

The nonresonant effects of higher J_{nm} coefficients are very small. In addition for m beyond 15 the orbit period for resonance is so small that air drag would be more serious than the resonance problem. For example, the 27th order harmonic in satellite 1963-49B has beat period of 5 days indicating near resonance but the effect is less than 30 meters in position. If the orbit period were closer to resonance the effect would be greater but the beat frequency would be long enough so that it could be treated as a linear rather than periodic effect.

$J_{13,13} = 0.52 \times 10^{-6}$ $\qquad \lambda_{13,13} = 10.4°$

$J_{15,14} = 0.08 \times 10^{-6}$ $\qquad \lambda_{15,14} = 19.6°$

$J_{14,14} = 0.56 \times 10^{-6}$ $\qquad \lambda_{14,14} = 15°$

32. COORDINATE SYSTEMS

There are a number of coordinate systems used in celestial mechanics and astronomy. These systems can be classified by three items.

(I) Center of Coordinates

1. Topocentric - origin at observation point on earth's surface.
2. Geocentric - origin at earth's center.
3. Heliocentric - origin at sun's center.
4. Selenocentric - origin at moon's center.
5. Planetocentric - origin at planet's center.
6. Barycentric - origin at center of mass.

(II) Plane of Reference

1. Horizon - tangent plane at observer's position.
2. Equator - earth's equatorial plane.
3. Ecliptic - mean plane of earth's motion about the sun.
4. Galactic - mean plane of the Milky Way galaxy.

(III) Reference Direction

1. Vernal Equinox (♈) - axis directed toward intersection of ecliptic and equatorial planes (at a specified date).
2. Greenwich Meridian
3. Polar Direction - axis of earth's rotation in northern direction.

As we sit at the earth's center of mass, if we could, the sun, in a course of a year would be seen to trace out a great circle path among the stars. This great circle is the ecliptic. If the earth's equator is traced out on the celestial sphere, this great circle will be the celestial (terrestrial) equator. These circles intersect with a dihedral angle of about

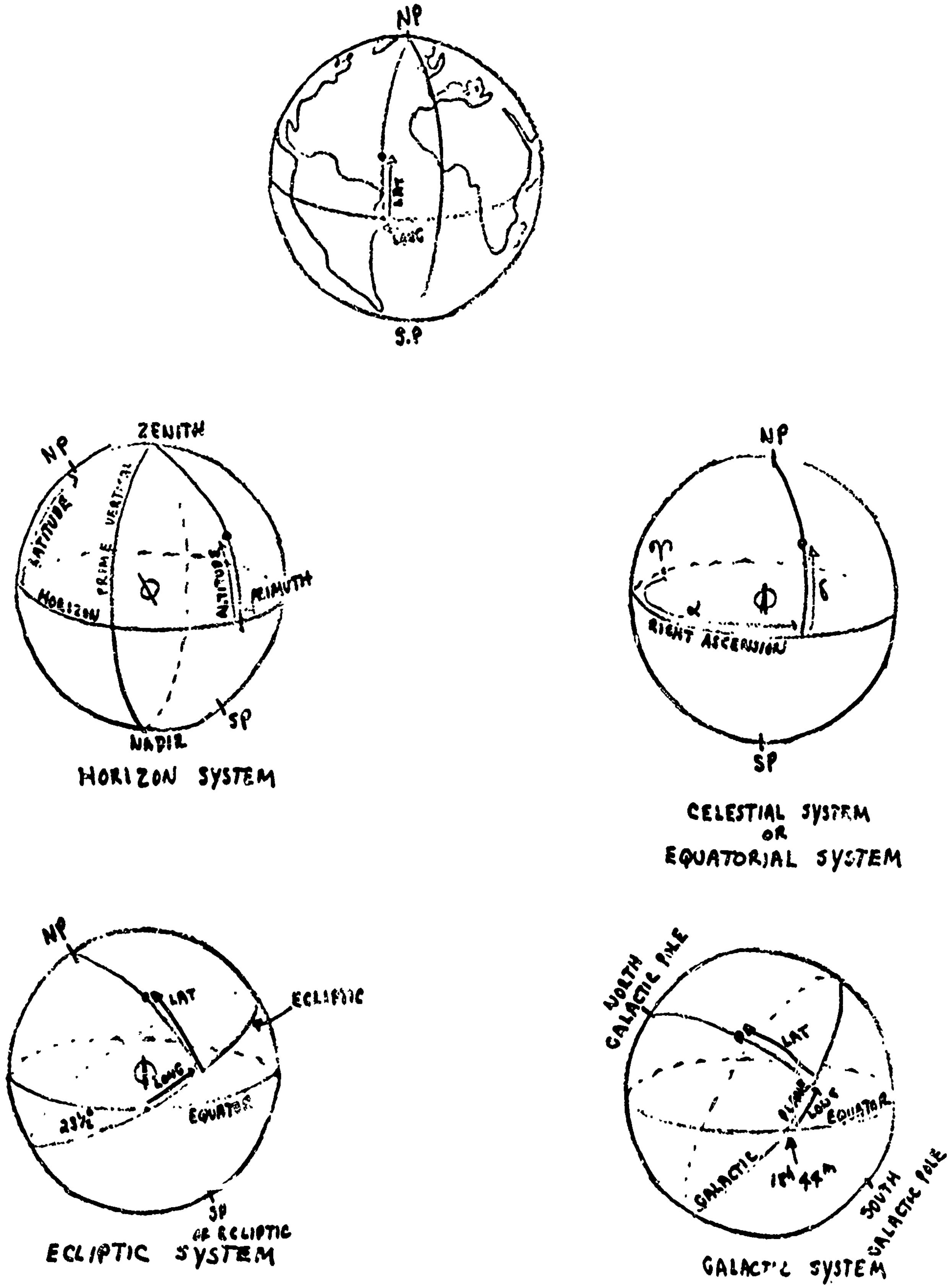


FIGURE 32-1 TERRESTIAL AND CELESTIAL COORDINATES.

23 1/2 degrees (23.44397° in 1964) and this angle, ε, is known as the obliquity. The vernal equinox, the point where the sun crosses the celestial (terrestrial) equator is used as a point of reference. This is the sun's position on about March 21 and is marked in the heavens by the so-called first point of Aries (♈).

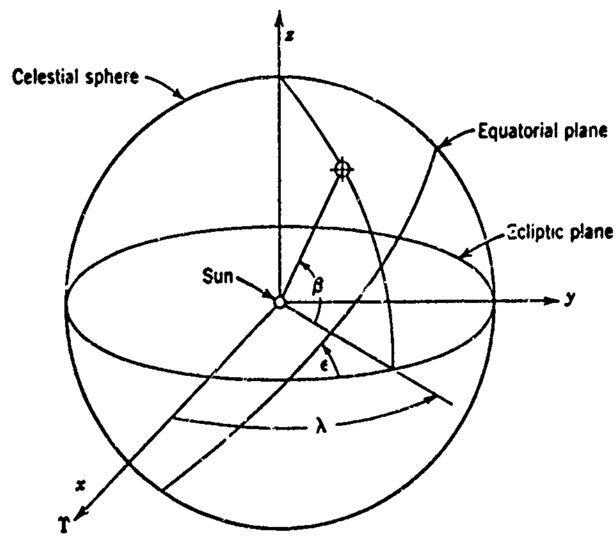


FIGURE 32-2. HELIOCENTRIC ECLIPTIC COORDINATE SYSTEM

The heliocentric ecliptic coordinate system has its coordinates centered in the sun, its plane of reference is the ecliptic and the x axis points to the vernal equinox. The y axis lies in the plane of the ecliptic, perpendicular to the x axis and the z axis points northward perpendicular to the x-y ecliptic plane. The angles needed to define a location in this system are the celestial latitude, β, which is measured normal to the ecliptic plane between a line connecting the object with the origin and

the ecliptic plane (see Figure 32-2); and the celestial longitude, λ, measured along the ecliptic from the x axis. β ranges from -90 to +90 degrees being positive above the ecliptic plane and λ ranges from 0 to 360 degrees, positive in a counterclockwise rotation as seen from above.

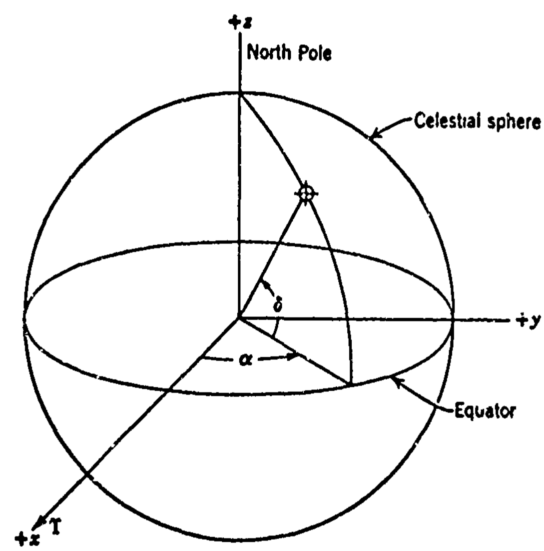


FIGURE 32-3. GEOCENTRIC EQUATORIAL COORDINATE SYSTEM

The geocentric equatorial coordinate system has its coordinates centered in the center of mass of the earth, its plane of reference in the equatorial plane and the x axis points toward the vernal equinox. Again the y axis is in the plane of the equator and the z axis goes through the north pole of the earth (see Figure 32-3). An object is located by α, the right ascension, which is measured in the plane of the equator from the fixed x axis, to a plane normal to the equator (meridian) which contains the object; and by δ, the declination, which is the angle between the object and the equator measured in a plane

normal to the equator as in Figure 32-3. α is measured in hours, minutes and seconds in the range 0 to 24 hours, positive counterclockwise as seen from the north pole. Thus 15° corresponds to one hour. δ is measured from -90 to +90 degrees, being positive above the equator.

One can also have a geocentric ecliptic coordinate system. Here one measures λ and β as in the heliocentric ecliptic system and these are related to α and δ through the obliquity by the following:

$$\begin{aligned} \cos\delta\cos\alpha &= \cos\beta\cos\lambda \\ \cos\delta\sin\alpha &= \cos\varepsilon\cos\beta\sin\lambda - \sin\varepsilon\sin\beta \\ \sin\delta &= \sin\varepsilon\cos\beta\sin\lambda + \cos\varepsilon\sin\beta \end{aligned} \tag{32-1}$$

and

$$\begin{aligned} \cos\beta\cos\lambda &= \cos\delta\cos\alpha \\ \cos\beta\sin\lambda &= \cos\varepsilon\cos\delta\sin\alpha + \sin\varepsilon\sin\delta \\ \sin\beta &= -\sin\varepsilon\cos\delta\sin\alpha + \cos\varepsilon\sin\delta \end{aligned} \tag{32-2}$$

Other coordinate systems are also used. A complete discussion and equations for transformation from one system to another are found in Chapter 4, page 125, of the book by Escobal, "Methods of Orbit Determination." See also Deutsch's book, Chapter 2, page 28, "Orbit Dynamics of Space Vehicles."

Note that there is a one-to-one correspondence between declination and geocentric latitude on the earth. The right ascension is related to the earth's longitude by the sidereal time of the Greenwich Meridian.

Using x, y, z as rectangular geocentric coordinates, see Figure 32-3, we have the relations,

$$\alpha = \tan^{-1}\frac{y}{x} \qquad \delta = \tan^{-1}\left[\frac{z}{\sqrt{x^2 + y^2}}\right]$$

$$\lambda_E = \text{geocentric earth's longitude} = \alpha - [\theta_g - \omega_E (t - t_g)]$$

Where the sign of y and x determine the quadrant for α and δ and θ_g is sidereal time of Greenwich Meridian (see Section 33), ω_E is earth's rotation rate and t_g is time of passage of the Greenwich Meridian and the x axis.

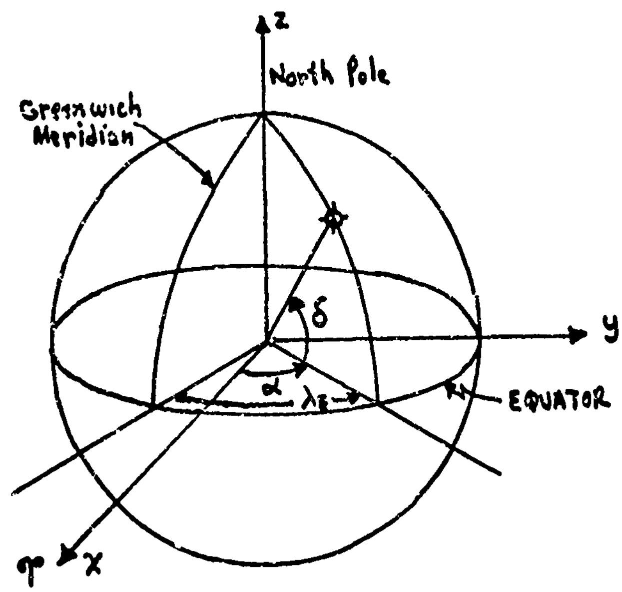


FIGURE 32-4. RELATION WITH GEOCENTRIC LONGITUDE

δ: declination = Geocentric Latitude.

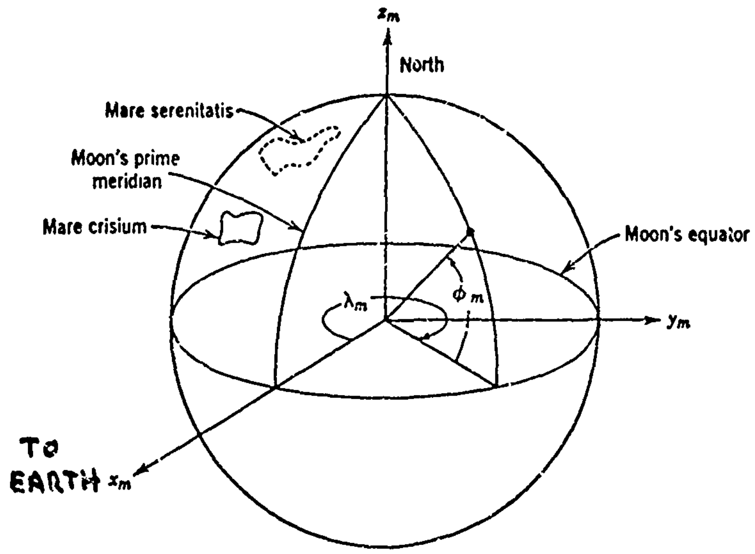


FIGURE 32-5. SELENOGRAPHIC COORDINATE SYSTEM

Now unfortunately for us, these systems so defined do not remain fixed in space as our axis system is required by Newtonian law to do. The sun and moon exert a righting force on the earth's equatorial bulge which attempts to pull the earth's equator into the plane of the sun and moon respectively. This force produces a gyroscopic turning force that is at right angles to both the axis of rotation and to itself, thus causing the axis of the earth, also our z axis, to sweep out a cone in space as shown in Figure 32-6. The period of this rotation is about 26,000 years so that the yearly motion is small indeed but it must be considered.

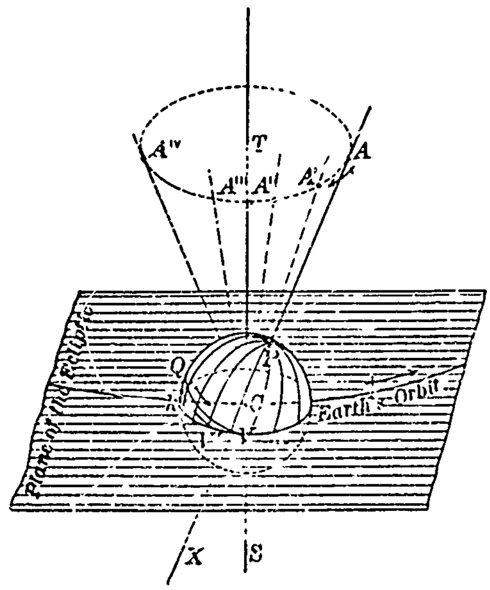


FIGURE 32-6. CONICAL MOTION OF EARTH'S AXIS DUE TO PRECESSION

The plane of the ecliptic was roughly described as that plane in which the earth moves around the sun, after allowances have been made for slight oscillatory motions of the earth's center, caused by the attraction of the moon and planets. Though small, being the order of tenths of a second of arc, these slight motions cause the apparent or true geocentric latitude of the sun to seldom be exactly zero. The attraction of the planets, besides contributing to these oscillations also cause the plane of the ecliptic to rotate nearly uniformly in space by about 47" per century. This righting effect tends to move the ecliptic plane into the invariable plane. This part of the precession is called planetary precession. It causes the equinox to move

eastward about 0.11" per year and it diminishes the obliquity by 0.47" annually. Therefore since the orbit of the earth about the sun is not truly planar, the "plane of the ecliptic" is not a precise notation - only a descriptive abstraction. The ecliptic one has in mind is essentially a mean which is defined by dynamical considerations.

Now the celestial equator is not the same sort of abstraction as the ecliptic. At any instant, the axis about which the earth instantaneously rotates, exists. It is this axis which defines the equatorial plane and is the axis which is precessed by luni-solar precession at a nearly uniform rate of 50" per year. This luni-solar precession combines with the planetary precession to produce what is called general precession.

All of these precessions cause the equinox to move along the ecliptic plane, as shown in Figure 32-7, with a westward motion. In addition, the equatorial plane is lowered, i.e., there is a change in the declination angle. The equinox is moved from ♈ to $♈_1$. All these motions represent a change in declination and right ascension in accordance with the equations below.

$$\frac{d\delta}{dt} = + 20''\!.0426 - 0.000085\ (T - 1950)\ \cos\alpha$$

$$\frac{d\alpha}{dt} = + 4''\!.40944 - 0.0000186\ (T - 1950) - 0.0000057\ (T - 1950)\ \sin\alpha\ \tan\delta$$

Fortunately these rates are so nearly uniform that one may compute the change in α or δ by multiplying the number of years in the interval

by the corresponding rate at the middle of the interval. Some values are shown below in seconds of arc per year.

	1959	1964
General Precession	50.2695	50.2706
Precession in α	46.1015	46.10295
Precession in δ	20.0398	20.0414

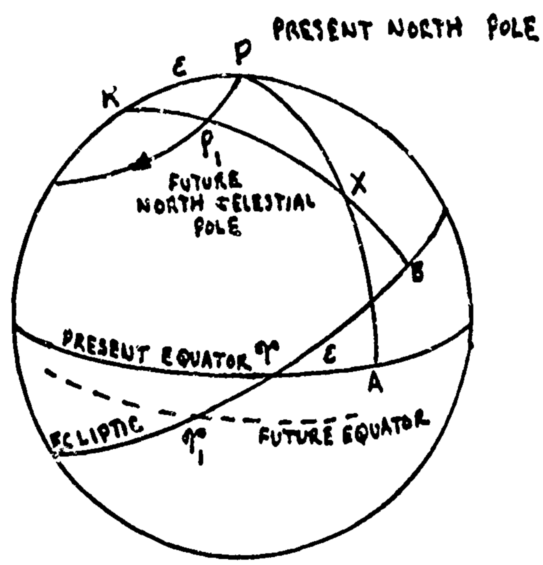


FIGURE 32-7. PRECESSION OF ECLIPTIC AND EQUATOR

These motions are called the mean motions. The actual celestial pole motions have an additional and periodic variations called nutations. When one considers the righting force of the sun and moon, it is seen that this force disappears whenever the sun or moon crosses the equatorial plane. This occurs twice a month for the moon and twice a year for the sun.

Hence the conical motions of the pole has superimposed upon it small wiggles called "nutations." These are never greater than 9.23" with a period of a little less than 19 years. These effect the equator location but not the ecliptic and thus alter celestial longitudes and the obliquity but not celestial latitudes.

It should be mentioned that these "nutations" having nothing to do with the conventional concept of nutations of a gyro precessing with pseudoregular precession. They are not the same phenomena and differ greatly in their periods.

Our coordinate system must be a fixed one, and is fixed to the equinox and equator, but we must decide which equinox and equator, since they vary. The instantaneous positions at any instant are called the mean equator and equinox of date. Thus the mean equator and equinox of an arbitrarily chosen epoch may be used to define a nonrotating ecliptic and equator. We may choose the equinox at time of launch of the missile or the mean equinox and equator at the beginning of 1950, one which is used for uniformity and convenience by most astronomers. For lunar trajectories the former is best, for interplanetary travel the latter is proper.

The Nautical Almanac and Ephermeris lists the coordinates of the sun and moon in a system referred to the mean equinox and equator at the beginning of the year and at the beginning of 1950. The moon's coordinates are also related to true equinox and equator of date.

It is necessary to resolve these coordinates into our fixed coordinate system by rotation through the precessed α and δ angular amounts.

REFERENCES

1. Escobal, P. R., "Methods of Orbit Determination," John Wiley, New York, 1965, Chapter 4 and Appendix I.
2. Roy, A. E., "The Foundation of Astrodynamics," Macmillan Company, New York, 1965, Chapters 2 and 3.
3. Russell, H. N., R. S. Dugan and J. Q. Stewart, "Astronomy," Ginn and Co., Boston, 1926.
4. Wolaver, L. E., "Factors in Precision Space Flight Trajectories," in WADC Technical Report 59-732, Vol. 1, Space Technology Lecture Series, P. 77, December 1958.

33. TIME

We will have our results in the form of position and velocity as a function of time. Let's examine just how one determines the time. The basic time keeper is the rotation rate of the earth.

One time unit is the apparent solar day which can be determined by the time between successive passages of the sun over the observer's meridian. This can be found by noting that the sun is on the meridian when the shadow cast by a vertical post is shortest. The problem with this is that the earth moves about the sun in an ellipse in which equal angles are not swept out by the radius in equal times and the ecliptic and equator are not in the same plane. This causes the length of day to vary. To overcome this, the length of a day is averaged over an entire year to form the mean solar day.

To be more precise, one assumes a dynamical mean sun which moves with the mean angular velocity of the sun (n_s) in the plane of the ecliptic so that the perigee-to-perigee time is the same as the actual sun. When this dynamical mean sun, moving in the ecliptic reaches the vernal equinox, a second fictitious body, the mean sun, moves along the equator with the sun's mean motion, returning to the vernal equinox with the dynamical mean sun. The time between successive passages of this mean sun over the observer's meridian is constant and defines the mean solar day.

A more precise standard can be had by considering the time between successive passages of a fixed star, or a particular fixed vernal equinox. This time is called one sidereal day.

Since the earth's radius vector sweeps out about 1 degree per day and the earth rotates at an angular velocity of about 1 degree in four minutes, the sidereal day is about 4 minutes shorter than the mean solar day.

1 mean solar day = $24^h03^m56.5554^s$ of sidereal time

1 sidereal day = $23^h56^m04.0905^s$ of mean solar time.

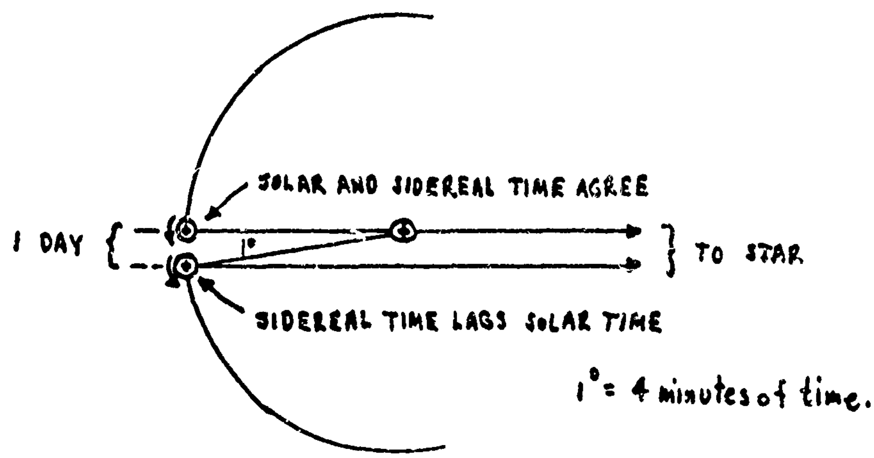


FIGURE 33-1. SOLAR AND SIDEREAL TIME RELATIONS

There are other problems because the earth's rotation is not precisely constant but before going into that let us consider the calendar problem.

The true length of the year is not 365 1/4 days but 365 days, 5 hours, 48 minutes, and 46.0 seconds leaving a difference of 11 minutes and 14 seconds by which the Julian year is too long. This Julian calendar was established by Caesar about 45 B.C. and adopted 365 day a year with every fourth year containing 366 days. This 11 minutes and 14 seconds amounts to little more than 3 days in 400 years. As a consequence, the Julian date of the vernal equinox came earlier and earlier and so Easter, defined as the first Sunday after the first full moon after the vernal equinox, began to arrive in the bare winter. To correct this unesthetic setting for Easter, Pope

Gregory in 1582 ordered the calendar corrected by dropping 10 days so that the day following October 4, 1582 was called the 15th instead of the fifth and in addition, thereafter only such century years should be leap years as are divisible by 400. Thus 1700, 1800, 1900, 2100 are not leap years while 1600 and 2000 are. England waited until 1752 and decreed the day following September 2, 1752 should be called the 14th instead of the 3rd, thus dropping eleven days. Even though the act was carefully framed to prevent any injustice in collection of interest, rents, etc., it nevertheless caused some people to riot, particularly in Bristol where several people were killed. The rally cry was, "Give us back our fortnight!"

Russian changed in 1918 and Rumania in 1919.

This gap in the calendar is a discontinuity. To avoid it, astronomers use the Julian year as proposed by J. Scaliger in 1582. It starts arbitrarily on January 1, 4713 B.C. and consists of 36525 days in each century. Thus the Gregorian date of January 24, 1925 is J.D. 2,424,175. A conversion table is given in the Nautical Almanac. Because astronomers work all night and don't want to keep changing the date in the log book, the Julian day begins at Gregorian noon. Thus when January 24, 1925 begins (0000 GMT), the Julian date is 2,424,174.5 and the date of an observation made on January 24, 1925 at the 18th hour Greenwich Mean Time is J.D. 2,425,175.25.

Orbital data is often referred to Modified Julian Day number in which the zero point is 17.0 November 1858 and hence

Modified Julian Date = Julian Date - 2,400,000.5 days.

In addition to the sidereal year, the time the sun returns to a fixed star ($365^D 6^H 9^M 9.5^S = 365.25636$) we also have the tropical year referred to above, which is the time between successive passages of the vernal equinox ($365^D 5^H 48^M 46.0^S = 365.24220$); and the anomalistic year, the time between successive perihelion ($365^D 6^H 13^M 53.0^S = 365.25964$ mean solar days). Because of precession the latter two are not constant values.

Now let us return to time. The Greenwich sidereal time is called Universal Time (U.T.) and both it and mean solar time are based on the rotation of the earth on its axis. This is not a precise mainspring. This fact was discovered when it was found that the position of the moon as calculated did not correspond to the observed position. The difference between these two values of longitude of the moon, for example, sometimes exceeds 10" of arc and may change by several seconds of arc in a few years. Similar but smaller differences have been observed in the longitudes of the sun, mercury and venus. They are all in the same direction in the same year and their amounts are proportional to the mean motions of these bodies. This indicates that the computed positions were correct, but they were computed for the wrong time. These may be explained as an altertation in the uniform rate of the earth's rotation.

One then proceeds to define what is called Ephemeris Time (E.T.) in order to have a uniform, continuously varying time scale as an independent variable for gravitational theory. This E.T. is defined so that the length of the day in Ephemeris Time is roughly equal to the average length of the day in Universal Time over the last three centuries.* The change from Universal to Ephemeris Time is accomplished by interpolating to a time ΔT seconds earlier when

(*To be precise there are 31 556 925.9747 ephemeris seconds in tropical year 1900.)

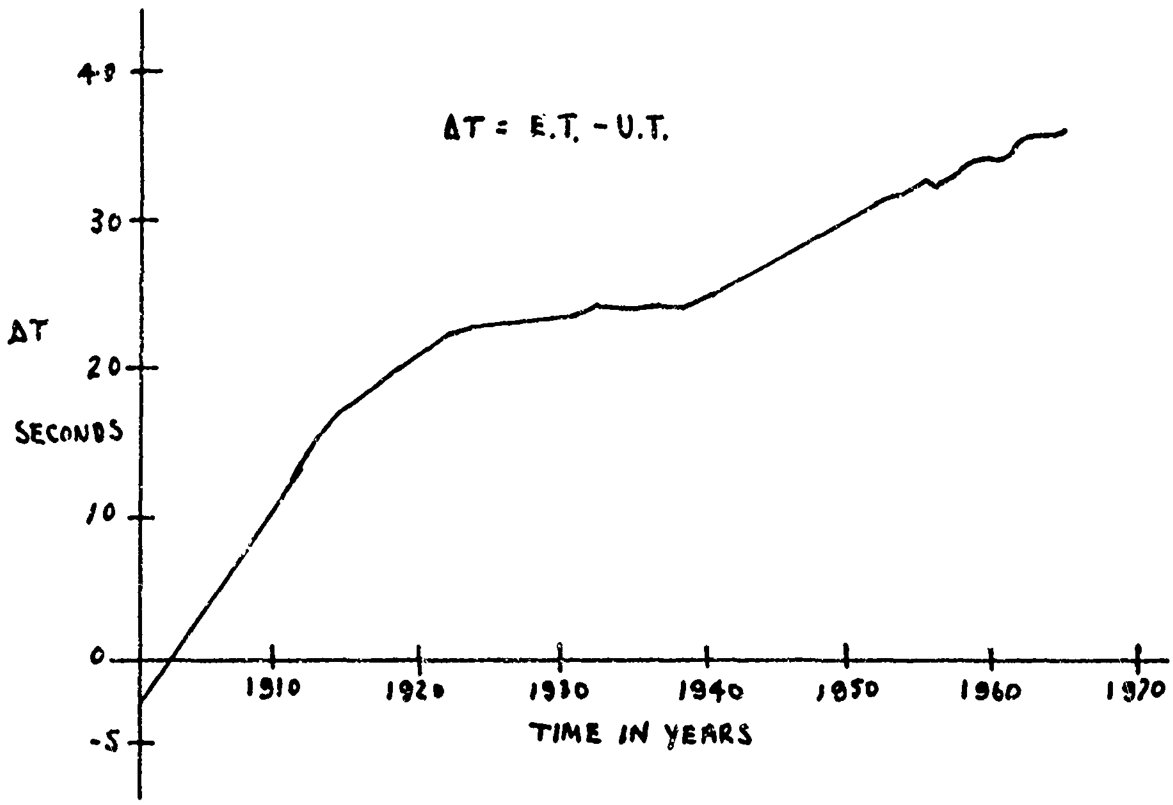
$$\Delta T = E.T. - U.T. = 24^{s}.349 + 72^{s}.3165T + 29^{s}.949T^2 + 1.8214B^* \quad (33\text{-}1)$$

Where T is reckoned in Julian centuries from 1900 January 0 Greenwich Mean Noon and B* is an emperical term derived from an attempt to reconcile observed and computed positions of the moon. The ΔT conversions are listed in the Nautical Almanac for each year. A few of these are listed below, see also Figure 33-2.

ΔT (seconds)	YEAR
+29.66	1951.5
+31.59	1955.5
+32.80	1958.5

This correction to our Greenwich Mean Time (U.T.) cannot be calculated in advance, only estimated. It is only after long reductions of the observed and predicted longitudes of the moon that B* and hence ΔT can be computed. Hence it is always currently an estimated quantity. It should be mentioned that this E.T. does indeed correspond to precise cesium clock time ($\pm$ 1 part in 10^{10}) and is not just a mathematical device.

This nonuniform rotation of the earth deserves a little more attention. In 1695, Halley using the results of ancient and recent eclipses was led to suspect that the mean motion of the moon was increasing (relative to the earth's rotation). Dunthorne in 1749 confirmed this and found that if a term $10''T^2$ (T in Julian centuries since 1900.0) were added to the calculated coordinates of the moon, it would correct the secular increase in the length of the day.


40
30
20
10
0
-5
ΔT = E.T. - U.T.
ΔT
SECONDS
1910
1920
1930
1940
1950
1960
1970
TIME IN YEARS

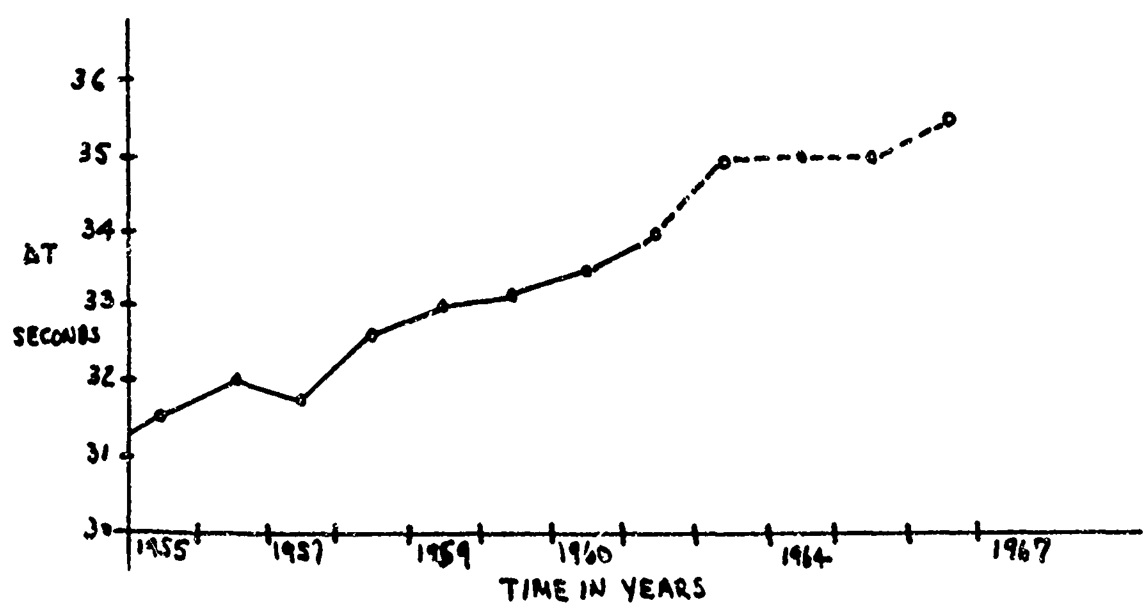

36
35
34
33
32
31
30
ΔT
SECONDS
1955
1957
1959
1960
1964
1967
TIME IN YEARS

FIGURE 33-2

If the moon's mean longitude is expressed as $a + bT + cT^2$, then the mean motion is $n = b + 2cT$ and the acceleration is $\dot{n} = 2c$. However, the coefficient c of T^2, not $2c$, is usually referred to as the acceleration.

Why this acceleration occurred remained a mystery until 1787 when Laplace announced the solution. Interestingly Euler and Lagrange tried but failed on this one. The explanation goes like this: The effect of the other planets upon the earth causes a gradual decrease in the eccentricity of its orbit. As the semi-major axis must remain constant, the semi-minor axis must be increasing, thereby increasing the mean distance of the sun, which in turn reduces its effect upon the moon - resulting in a slow increase in the moon's mean motion. Laplace's value of c was 10".18 which agreed nicely with the observed one. However he used a series solution, neglecting what he supposed were negligible higher order terms. In 1853, J. C. Adams found that several of the neglected terms, though small, actually added in phase to produce a measurable effect and he revised the calculated acceleration value to 5.70". The problem then was to determine what caused the 4.3" difference. Brown in 1897 and 1908 obtained c = 5.82" and Spencer Jones in 1932 and 1938 obtained c = 5.22".

The 4.3 seconds difference is in part, at least, probably due to the effect of tidal friction, particularly in shallow seas. However the paucity of data makes it difficult to estimate with higher precision. The rate of increase in the length of day needed to account for this 4.3" secular acceleration of the moon requires a dissipation of more than 2100 million horsepower of tidal friction. If the depth of water and current

velocities are accurately known, the dissipation of energy in any sea area could be calculated. At first it was found that neither viscosity nor turbulence in the open ocean could cause sufficient dissipation. Then in 1920, G. I. Taylor pointed out that since the rate of dissipation is proportional to the cube of tidal current in shallow and turbulent flow, the greatest amount of dissipation would come from the long bays and channels where the tide rise and fall is large. He found the Irish Sea contributed 2% of the required amount. Jeffreys in the same year showed the Bering Sea contributed 66-80% of that required to explain the acceleration. However, Munk and MacDonald in their book, "The Rotation of the Earth," comment that, "...there is sufficient accord... that the problem has been considered as solved for 40 years. We wish to reopen it." They dispute the calculations of Jeffreys and show they fall short of the required 2100 million horsepower dissipation needed. The whole problem is masterfully discussed in their Chapter 11, page 175.

We can presently say that tidal friction plays an important part in the acceleration term c, but we do not have enough data to confirm its precise role. The discovery of deep counter currents which strike the continental shelf could also be a contributor. But in addition to tidal friction there is more.

The eccentricity of the earth's orbit changes with a period of about 24,000 years. However for any discussion over a period of several thousand years it is more conveniently listed as a secular term as $e = e_0 + \alpha t$. Similarly the measured discrepancies of the longitude of the moon (difference between observed and calculated results) can be represented by

$$\Delta L_{☾} = a_{☾} + b_{☾}T + f_{☾}(T) \tag{33-2}$$

Where $a_{☾}$ and $b_{☾}$ are empirically determined constants. The discrepancy $f_{☾}(T)$ is plotted in Figure 33-3.

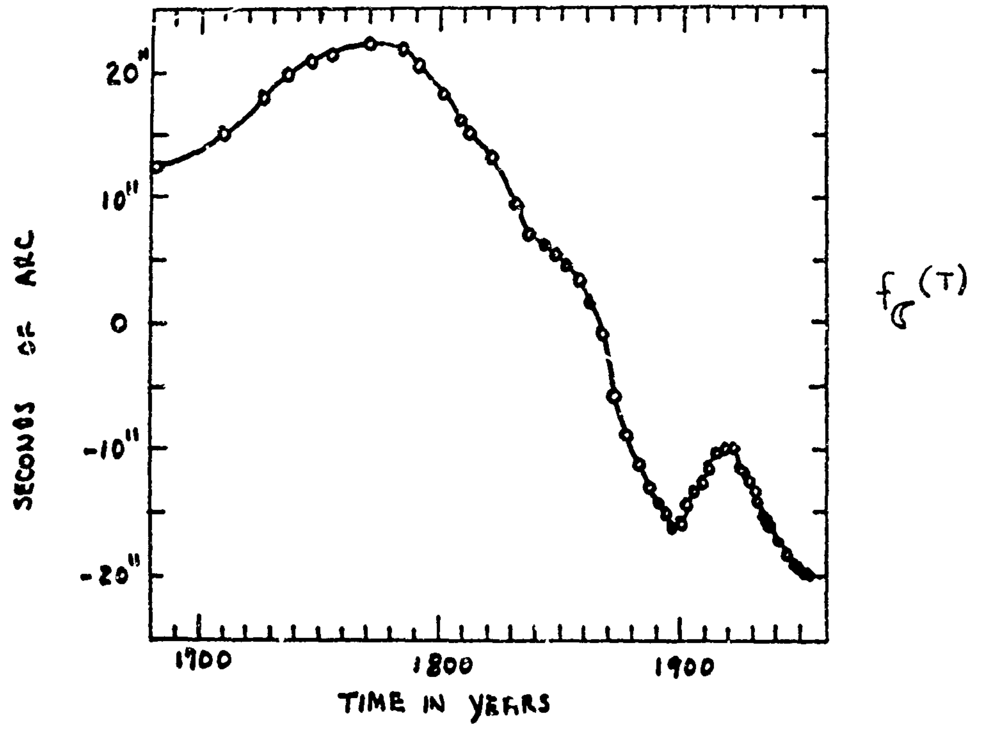


FIGURE 33-3. THE DISCREPANCY $f_{☾}(T)$ OF THE MOON'S LONGITUDE BASED ON OCCULTATIONS.

Three features should be noted: (1) a "secular" decrease; (2) a smooth hump between 1680 and 1850 which represents Newcomb's "Great Empirical Term" (G.E.T.); and (3) a relatively high frequency wiggle from 1900 to 1950, called the twentieth century wiggle.

Figure 33-4 shows this same longitude discrepancy for the sun and mercury, multiplied by the ratios

$$\frac{n_{☾}}{n_{\odot}} \approx 13.37 \qquad \frac{n_{☾}}{n_{☿}} = 3.22$$

If the variable rotation of the earth were the sole cause of the discrepancies, then a discrepancy, ΔT, between U.T. and E.T. would produce a discrepancy in longitude, ΔL, which is proportional to the mean motion of the bodies as observed from earth. Under these assumptions the discrepancies of moon, sun and mercury as in Figures 33-3 and 33-4 should be the same. There is, in fact, a close resemblance but it is not complete. There is a long term drift between lunar discrepancies (33-3) and those of mercury and the sun (33-4).

To see the difference we form the weighted discrepancy difference WDD $= f_{☾}(T) - \frac{n_{☾}}{n_{\odot}} f_{\odot}(T)$ and also between the moon and mercury. When we do this the WDD for both the sun and mercury can be represented reasonably well by the parabola, WDD $= - 11.2T^2$ seconds of arc. See Figure 33-5.

The WDD curve shows no sign of the Great Empirical Term or the twentieth century wiggle, so these fluctuations are common to the moon, sun and mercury and hence arise from variations in the earth's rate of rotation.

If we assume that, aside from a known quadratic term, the discrepancy in longitude of the sun is precisely equal to that of the moon times the ratio of their mean motions, then we can translate lunar observations into differences between Ephemeris and Universal Time.

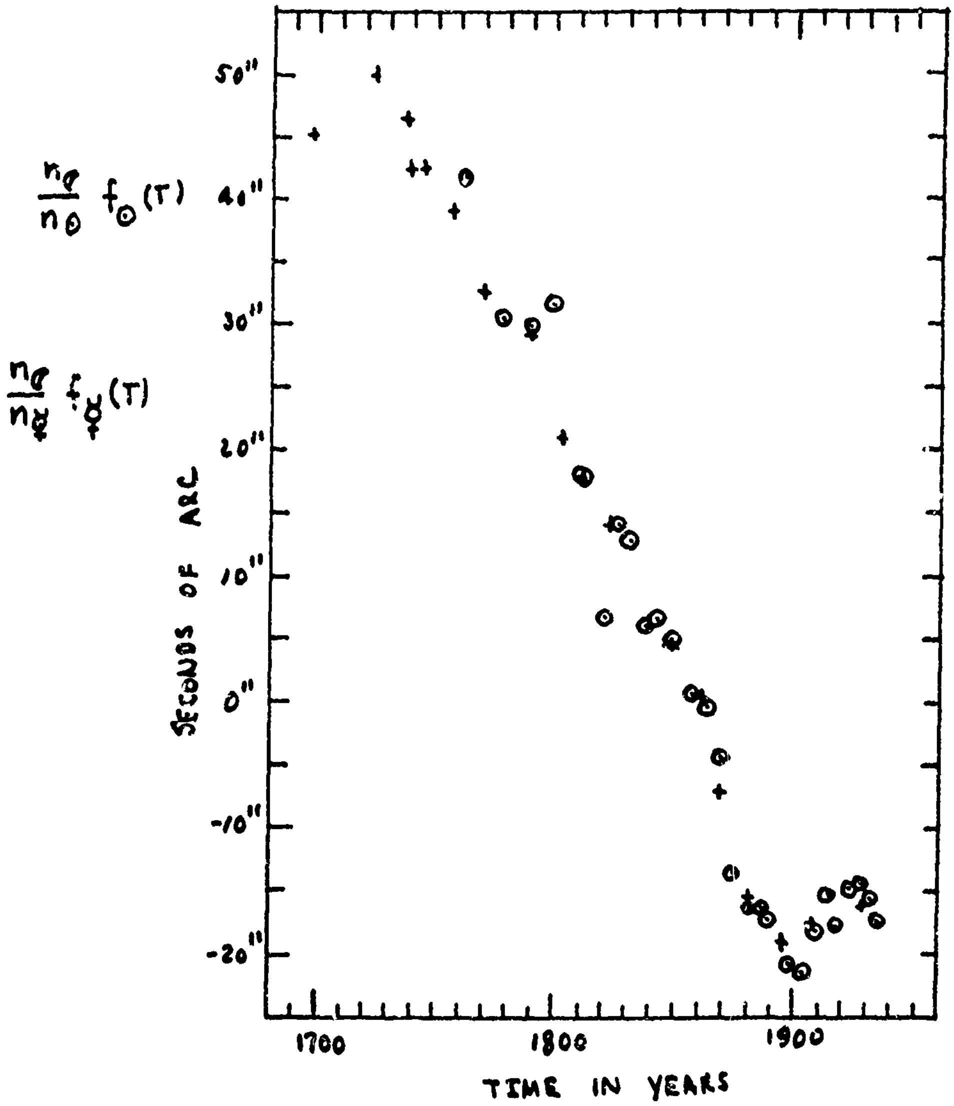


FIGURE 33-4. DISCREPANCIES OF SUN AND MERCURY WEIGHTED FOR THEIR MEAN MOTIONS. SEE "THE ROTATION OF THE EARTH."

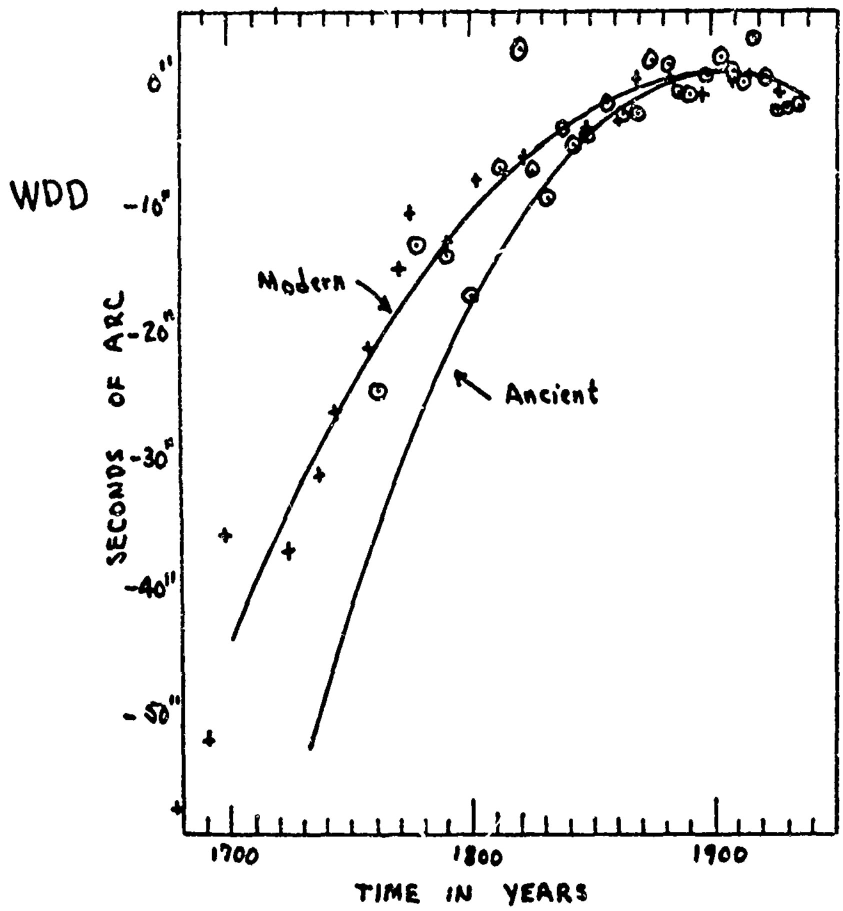


FIGURE 33-5. WEIGHTED DISCREPANCY DIFFERENCES ⊙ FOR THE SUN; + FOR MERCURY. CURVE LABELED "MODERN" CORRESPONDS TO WDD = $-11''.2T^2$; "ANCIENT" TO $-18''.85T^2$. SEE "ROTATION OF THE EARTH."

$$\Delta T = ET_{\odot} - UT = 24.349 \left[a_{\odot} + b_{\odot}T + \frac{n_{\odot}}{n_{☾}} cT^2 + \frac{n_{\odot}}{n_{☾}} f_{☾}(T) \right] \qquad (33\text{-}3)$$

Where T is in ET units. Applying the proper data gives the numerical values of equation (33-1).

Besides this secular acceleration there is a seasonal change in rotation rate of the earth which is included in the $f_{☾}(T)$. These seasonal changes were detected by quartz crystal clocks in 1935. The total variation in the length of the day during the year is about 2 1/2 milliseconds. The effect varies but the general trend is to cause the earth to be slow in spring and fast in autumn relative to a uniform time.

Seasonal fluctuations such as shifts in air masses, melting of the polar ice-cap, variations in the angular momentum of the atmosphere could be expected to cause these results, but there is too little data to compute precisely.

There are other fluctuations which appear suddenly and are of the order of as much as 5 milliseconds per year. In particular the 1958 anomaly seems well established from cesium clock measurements. See Figure 33-7. The cause is unknown. Perhaps an electromagnetic coupling of the mantle to a possible turbulent core could cause such disturbances. This would require the conductivity of the mantle to be about 10^{-9} e.m.u. which is high but not entirely outside the possible limits. We shall chalk this one up to the twisting churning insides of a dissipated world.

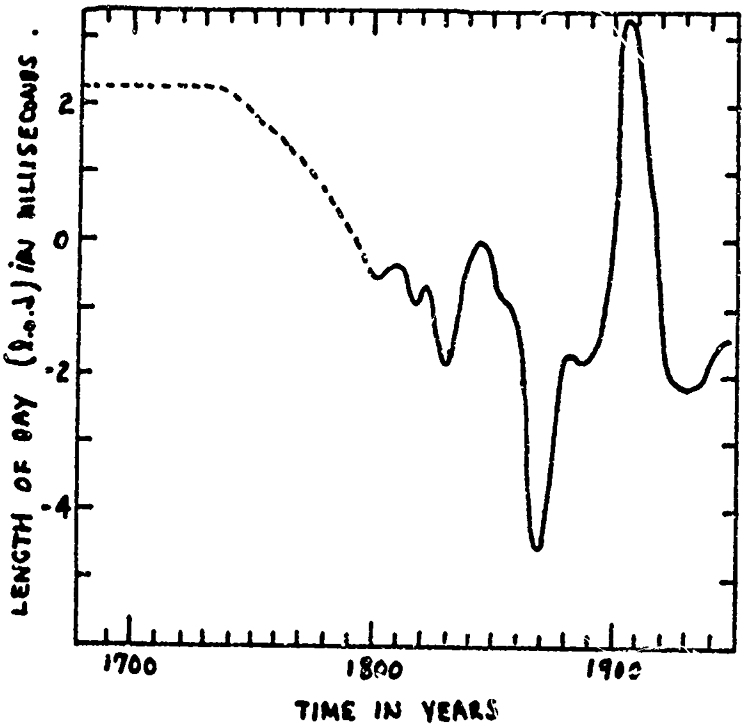
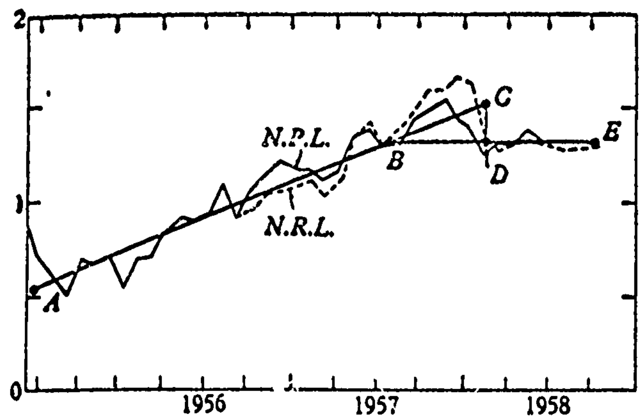


FIGURE 33-6. NONTIDAL VARIATION IN THE LENGTH OF THE DAY. PRIOR TO 1800 A SMOOTHED CURVE IS DRAWN.



The length of day derived from PZT observations (Washington and Richmond combined), referred to the caesium standards at the National Physical Laboratory, Teddington, and at the Naval Research Laboratory, Washington, D.C. Values for 1955, 1956 and 1957 are from Essen, Parry, Markowitz and Hall (1958); subsequent values have been kindly transmitted to us by Markowitz. Ordinate is milliseconds in excess over 86,400 seconds of Ephemeris Time

FIGURE 33-7. LENGTH OF DAY DERIVED FROM CESIUM STANDARDS AT NATIONAL PHYSICAL LABORATORY AND NAVAL RESEARCH LABORATORY. ORDINATE IS MILLI-SECONDS IN EXCESS OVER 86,400 SECONDS OF EPHEMERIS TIME.

REFERENCES

1. Munk, W. H. and G. J. F. MacDonald, "The Rotation of the Earth - A Geophysical Discussion," Cambridge Press, London, 1960.

2. Spencer Jones, H., "Dimensions and Rotation," Chapter 1 in the Book, The Solar System, Vol. II, The Earth as a Planet," edited by G. P. Kuiper, University of Chicago Press, Chicago, 1954.

3. Wolaver, L. E., "Factors in Precision Space Flight Trajectories" in WADC Technical Report 59-732, Vol. 1, Space Techrology Lecture Series, P. 77, 1958.

34. PARAMETRIC VALUES

In solving problems numerically one must eventually grapple with the problem of what numbers to use to represent the various constants. The first detailed discussion on the evaluation of astrodynamical constants of recent times was by de Sitter.[5] The various interrelations between constants and their best presatellite values is masterfully discussed by Herrick, et.al.[9] Other recent discussions are given in the references by Mercer,[20] Clarke[4] and by Brouwer.[3]

Consider the gravitational parameter $\mu = k^2 M_e$, where k^2 is Gauss' constant and M_e is the mass of the earth. One way to calculate this constant is to determine γ and M_e where $\mu = k^2 M_e = \gamma M_e$ and γ is Newton's universal gravitational constant. The International Critical Tables (1926) give

$$\gamma = 6.658 \times 10^8 \ Cm^3/g - sec^2$$

$$M_e = 5.988 \times 10^{27} \ grams$$

$$\gamma M_e = 3.9868 \times 10^{20} \ Cm^3/sec^2$$

These values are not known very accurately and a better scheme is to use the methods of terrestrial geodesy and gravimetry to define an equipotential surface called the geoid. This can then be used to find an equivalent spherical mass. The principal source of error here is in the triangulation and gravimetry measurements and the fact that so much of the world lies unmeasured. Kaula (1961) extrapolated from observed gravity measurements to obtain estimates for areas with gravimetry by

assuming the gravity anomalies could be treated as a Markov-Array, i.e., that the most probable value for an area without observations could be treated as a function of only the nearest observations. It is expected that the μ value so obtained will be smaller in absolute magnitude than the true value.

Uotela (1962) determined the geoid by using harmonic coefficients to fit the mean gravity anomalies of 5° x 5° squares by simple least square fit.

His coefficients tend to be larger in absolute magnitude than the true values. In any event such geoid determinations are not as accurate as satellite observations because of the paucity of gravitmetric data, especially over water. Some recent geodetic determinations give -

Fischer[7] (1962)	$3.986040 \times 10^{14}\ m^3/sec^2$
Kaula[20] (1961)	$3.986020 \times 10^{14}\ m^3/sec^2$

If one corrects Kepler's third law for variations of the lunar orbit due to the sun and other planets and to the shape of the earth, one can write (see Equation (23-28), page 231)

$$n_{\mathbb{C}}^2 a_{\mathbb{C}}^3 = 0.99728189\mu \left[1 + \frac{M_m}{M_e}\right] \qquad (34\text{-}1)$$

Then by measuring $a_{\mathbb{C}}$ by radar or by triangulation we can determine μ. Using $\frac{M_m}{M_e} = 81.301$, Yaplee, et al.,[21] calculated the μ value using radar

measurements over a long period of time. O'Keefe and Anderson determined the $a_{☾}$ value by triangulation. The principal errors are the uncertainty in the effective radar lunar radius. The constants were found to be

Yaplee, et. al.[21] 1963	3.986057×10^{14} m^3/sec^2
O'Keefe and Anderson[20]	3.986057×10^{14} m^3/sec^2

It is encouraging that all these values are so close. The most accurate determination is by satellite measurements. Some of these are listed below:

	10^{14} m^3/sec^2
Best Presatellite (Herrick)[9]	3.9861794
Echo I	3.986037 ± 12
Early Vanguard Data (O'Keefe)	3.98618
Later Vanguard Data (O'Keefe)	3.986062 ± 145
Lunar Probe and Doppler (Anderson, et.al.,1964)[1]	3.986032 ± 30
Close Satellite and Camera (Kaula 1964)[20]	3.986015

NASA[4] and Aerospace Corporation[20] have standardized on = 3.986032 ± 32. JPL uses a value of 3.986012. The International Astronomical Union (IAU) has adopted 3.986030×10^{14} m^3/sec^2 ± 0.00002. In all of these recent determinations it is the sixth decimal figure that differs. Thus one is always safe in using 3.9860×10^{14} m^3/sec^2. The NASA standard value[4] is

$$\mu = 3.986032 \times 10^{14} \text{ m}^3/\text{sec}^2$$

$$= 1.407653916 \times 10^{16} \text{ ft}^3/\text{sec}^2 \qquad (34\text{-}2)$$

$$= 0.5530393477 \times 10^{-2} \text{ (earth radii)}^3/\text{min}^2$$

$$= 1.536223188 \times 10^{-6} \text{ (earth radii)}^3/\text{sec}^2 .$$

Another constant we require is the earth to sun mass ratio. This is determined by developing a relation between the solar parallax or astronomical unit (A) and the mass ratio. The astronomical unit (A) is defined as (see page 10)

$$A^3 = 2.5226941 \times 10^{13} \, k^2 M_s \qquad (34\text{-}3)$$

This together with equation (34-1) which can be written

$$a_{☾}^3 = 1.4076669 \times 10^{11} \, k^2 (M_e + M_m) ,$$

One can obtain the relation,

$$\frac{A}{a_{☾}} = 5.6379545 \sqrt[3]{\frac{M_s}{M_e + M_m}} \qquad (34\text{-}4)$$

The numerical factor on the right side depends on observed data through the mean motion of the sun and moon only - and since time is

one thing we can measure accurately - it is correct to the last place given. Then making use of

$$\frac{a_☾}{a_e} = \frac{3.844002 \times 10^8 \ (1 \pm .05 \times 10^{-5})}{6.378166 \times 10^6 \ (1 \pm 0.40 \times 10^{-5})}$$

$$= 60.26814 \pm 0.00039 \qquad (34\text{-}5)$$

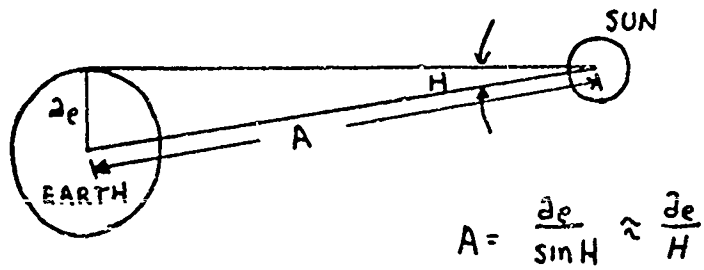
we can write

$$\frac{A}{a_e} = (339.7890 \pm 0.0022) \sqrt[3]{\frac{M_s}{M_e + M_m}} \qquad (34\text{-}6)$$

Thus if we measure A we have effectively computed the mass ratio. Rather than use A one measures a quantity called the solar parallax. This is the sun's mean equatorial parallax measured in seconds of arc. It appears in this form in the reduction of observations from geocentric to topocentric* positions and is therefore of fundamental interest to astronomers. If we call H the solar parallax then

$$\sin H = \frac{a_e}{A} \qquad (34\text{-}7)$$

(* see page 398)



Making use of Equation (34-i) we have (*H in seconds of arc.*)

$$H \sqrt[3]{\frac{M_s}{M_e + M_m}} = 607.0379 \pm 0.0039$$

or

$$H^3 \left(\frac{M_s}{M_e + M_m} \right) = 2.236904 \times 10^8 \pm 43 \quad . \tag{34-8}$$

This relation between the solar parallax and mass ratio is more accurate than previous relations based on Equation (34-i) together with a geophysical development for $\mu = k^2 M_e$. For example see the derivation given by de Sitter[5] where he finds

$$H^3 \frac{M_s}{M_e + M_m} = 2.2346135 \times 10^8 \pm 140 \tag{34-9}$$

Brouwer[3] suspects this latter relation is more liable to have systematic errors and I am not one to suspect Brouwer. Howsoever we have

$$\frac{M_s}{M_e + M_m} = \frac{2.236904 \times 10^8}{H^3} \qquad (34\text{-}10)$$

so that a measurement of H will determine the mass ratio. One method is to determine the solar parallax by observation of the angular displacement of an object in the solar system with respect to the background of fixed stars when observed from two stations separated a known distance on the earth's surface. The distance to the object is then found by triangulation. The most accurate determinations are made on minor planets which come close to the earth and therefore have a large parallax angle. Minor Planet No. 433 (EROS) comes within 1/5A (about 28×10^6 km) of earth and has a parallax of 50" which is 5.7 times greater than the solar parallax. Its close approaches to earth during 1930-1931 together with previous photographic data was reduced by Spencer Jones.[10] He took ten years, worming over 2847 photographic plates taken with 30 different telescopes at 24 observations in 14 countries. He arrived at the value $H = 8''.7904 \pm 0.0015$ (mean error). Using Equation (10) this corresponds to a mass ratio of 329323 ± 170 and $A = \frac{a_e}{H} =$ 149,662,400 km $\pm$ 256 km.

A second and more confidence inspiring method is the dynamical method used by Rabe.[16] Here one determines the mass ratio directly

from the perturbations caused by the earth-moon system on the asteroid Eros. It requires the calculation of a precise orbit of Eros over a period of several revolutions. The correction to the mass of the earth-moon used in these calculations is one of a large number of unknowns which include the other planet masses. The calculations are long and intricate and the unknowns are evaluated by using least square fit to the observed motion. Eugene Rabe used Eros and took into account the perturbations by Mercury, Venus, Earth-Moon, Mars, Jupiter, Saturn and Neptune. He derived a value of

$$\frac{M_s}{M_e + M_m} = 328452 \pm 67 \, . \qquad (34\text{-}11)$$

This corresponds to a solar parallax of $8''79816 \pm 0.0006$ and $A = 149{,}530{,}300 \pm 10{,}200$ km.

This difference between Rabe and Spencer Jones is several times the combined probable error of each and was the subject of much discussion.

A third method to determine A directly is available with radar. Since the solar system is well mapped by angular measurements, the determination of one distance leg will suffice to determine A. Hence a measurement of Earth to Venus distance by means of radar will also determine A. To calculate A by radar we must know the exact distance of the radar site from the center of the earth, the diameter of the planet Venus and a precise knowledge of the speed of light.

The most recent determination[6] gives the velocity as $c = 2.997925 \times 10^8$ m/sec $\pm$ 3. Some radar determinations of the astronomical unit are listed below:

A(km) (for c = 299792.5 km/sec)	
149,599,500 $\pm$ 800	USSR Koletnikow 1961[11]
149,601,000 $\pm$ 500	Thomson, et. al. Jodrell Banks 1961[18]
149,597,850 $\pm$ 400	Pettengill, et. al. Millstone Hill 1962[15]
149,598,595 $\pm$ 500	Muhleman, et. al. 1962[14]
149,599,400	JPL (1963) See Marsden (1963)[13]

This value of A = 149,599,400 km corresponds to $H = 8''.79413$ and a mass ratio of 328905.8. At this point in time there were three separate mass ratios. The difference in the dynamical and radar determinations was an important point which received increased attention. The radar people began a very careful error analysis of their work leading to a much greater understanding of the problem than would have developed if these three determinations were more similar. Finally in 1967 Rabe and Mary P. Francis[28] recalculated a mass ratio which invalidated Rabe's previous calculations. They obtained the mass ratio of 328899. $\pm$ 15. More recent radar determinations substantiated this figure. For example, Ash, Shapiro and Smith[22] (1967) obtained a mass ratio of 328900. $\pm$ 60 and Sjogren, et. al.[23] (1966) obtained 328900. $\pm$ 1. Thus there is no longer any discordance between radar and dynamical determinations of the Earth plus Moon to Sun mass ratio. Once again Newton's law of gravity has met the test and passed.

The principal determinations are summarized below:

	H seconds of arc	$\frac{M_s}{M_e + M_m}$	A(km)
Spencer Jones	8.7904 $\pm$ .0015	329323 $\pm$ 170	149,662,400
Rabe (1954)	8.79816 $\pm$ .0006	328452 $\pm$ 67	149,530,300
Rabe & Francis (1967)	8.79416 $\pm$ .00013	328899 $\pm$ 15	149,598,300
Radar (1961-1963)	8.79413 $\pm$ .000018	328905.8 $\pm$ 2	149,599,400
Radar (1966-1967)	8.79416 $\pm$ .000009	328900 $\pm$ 1	149,598,388 $\pm$ 50

Thus the most likely value is $\frac{M_s}{M_e + M_m} = 328900$, which corresponds to H = 8.79416 seconds of arc and A = 149,598,388. Using $\frac{M_e}{M_m} = 81.301$ this mass ratio gives

$$\frac{M_s}{M_e} = 332{,}945 \pm 1 .$$

NASA recommends[4]

A = 149,549,000 $\pm$ 700 km.

A = 0.4908103674 x 10^{12} feet

A = 23454.86515 earth radii.

Other parameters we need are the equatorial radius of the earth and the eight other planetary masses. Some recent values for the equatorial radius are listed below

Presatellite (Herrick, et. al. 1958)[9]	6378145 Meters $\pm$ 70
Moon Radar (Yaplee, et. al. 1958)[21]	6378175 Meters $\pm$ 35
Kaula (1961) geodetic	6378163 Meters $\pm$ 21
NASA Adopted (Clarke 1964)[4]	6378165 Meters $\pm$ 25.

The equatorial radius defined here is the mean equatorial radius.

If the equatorial plane were truly an ellipse in shape this would be the semi-major axis.

Another parameter is the mean distance between the center of the earth and the center of the moon, $a_{☾}$.

This has been accurately determined by a long series of radar measurements by the Naval Research Laboratory.

$$a_{☾} = 384{,}400.2 \pm 1.1 \text{ km}$$ (Yaplee, et.al.[21]).

Another item we need is the moon-to-earth mass ratio. For lunar trajectories this is a very important quantity.

The earth's axis moves in a cone about an axis perpendicular to the plane of the ecliptic. This average motion is called the precession and small harmonic motions superimposed upon it are called nutations. These are due to the attraction of the sun and moon on the earth's equatorial bulge. These precession and nutation factors are both functions of a moment ratio $\frac{C-A}{C}$ and the mass of the moon. Hence observations of the two phenomena provide two equations from which we may eliminate the moment ratio to compute the mass of the moon.

A more direct method is now employed by utilizing the perturbations of the earth's center caused by the moon. Since the earth's center moves in an ellipse about the center of mass of the earth and moon together, the direction of an external body as seen from the earth will suffer a monthly perturbation having an amplitude dependent upon the mass of the moon. Observations of the sun were employed by Newcomb

where he found the amplitude of the sun to be about 6 seconds of arc. For a close approach of an asteroid, however, the amplitude can exceed this magnitude several times.

The first use of an asteroid was by Gill on the Asteroid Victoria. Hink's study was based on the 1900-1901 opposition of Eros; H. Spencer Jones' (1942) study was made of the 1930-1931 opposition of Eros. Rabe used his dynamical equations to compute the mass ratio. Some of the values so determined are listed below:

Newcomb Observation of Sun	81.215
Gill on Asteroid Victoria 1888-1889	81.702 $\pm$ 0.094
Hinks on Asteroid Eros 1900-1901	81.53 $\pm$ 0.047
Spencer Jones on 1930-1931 Eros	81.271 $\pm$ 0.021
Rabe (1950) Eros	81.375 $\pm$ 0.028
Nautical Almanac and Ephemeris	81.45 .

The latter value was arrived at by a committee for standardization purposes and is only faintly related to the actual value.

Prior to sending satellites to the moon, Rabe's figure of 81.375 was widely accepted as the most reliable one. However from radio and radar tracking of Mariner to Venus and from the Ranger Lunar mission, the following mass ratios were found.

Earth to Moon mass ratio

81.3030 $\pm$ 0.0050	Ash, Shapiro, Smith 1967[22]
81.3001 $\pm$ 0.0013	Anderson 1967[1]
81.3020 $\pm$ 0.0017	Sjogren et.al. 1966[23]
81.3015 $\pm$ 0.0016	Null, Gordon and Tito 1967[24]

The value of 81.301 is currently widely accepted as the most probable one. Lunar satellities will improve this figure and also increase our knowledge of the shape of the moon which is very much a triaxial ellipsoid. Some recent references on the determination of the shape of the moon are those of Goudas and Kopal.[27]

The remaining values of k^2Mp for any planet, p, can be obtained from the relation

$$k^2Mp = k^2Me \frac{M_p}{M_s} \frac{M_s}{M_e}.$$

Thus the planetary ratios are all that are required.

To determine the mass ratios one uses a satellite of the planet. If the disturbing effects are ignored, Kepler's third law for satellite motion may be written

$$\frac{M_p + M_s}{M_\odot} = a^3 \left(\frac{n}{R}\right)^2$$

The observed sidereal motion, n_{obs}, is defined as the coefficient of t in the linear function of the time that results if the true longitude derived from the observations is freed from all the periodic terms and the precession. The constant a_{obs} is obtained by subtracting from the radius vector the elliptic terms and the periodic perturbations. Both n_{obs} and a_{obs} so obtained contain contributions due to the presence of disturbing forces. Allowance is made for these effects by modifying the formula expressing Kepler's third law to

$$\frac{M_p + M_s}{M_{\odot}} = a^3_{obs} \left(\frac{n_{obs}}{k}\right)^2 \left[1 - \frac{\sigma}{2}\right]$$

where σ is an analytic expression for the secular perturbation (see Reference 2, page 55 for further details). σ is a very small quantity so its square has been neglected in the modified formula.

Thus Kepler's third law with its appropriate correction factor is available for evaluating the masses in terms of the sun's mass. The only exception is the mass of the earth-moon since the moon's mean distance in astronomical units cannot be obtained from observations of the moon alone. For all other satellities the necessary information concerning their orbits is derived from measurements of the position of the satellite relative to the center of the disk of the primary. For planets with no satellities one relies on a dynamical method such as that used by Rabe. Unfortunately such calculations are ill-conditioned. In particular the use of Eros' orbit is apparently incapable of yielding reliable masses for any planet other than the earth.

A list of some planetary mass ratios are given on the next page.

Planetary Mass Ratios

Mercury	Brouwer 1950	6,480,000
	Rabe 1954	6,118,000
	Duncombe 1958	5,970,000 ± 455,000
	Ash, Shapiro and Smith 1967	6,021,000 ± 53,000
	I.A.U. (1966)	6,000,000
	J.P.L. (1969)	5,983,000 ± 25,000
Venus	Clemence 1943	409,300 ± 1400
	Rabe 1954	408,651 ± 208
	Anderson 1967	408,505 ± 6
	Ash, Shapiro and Smith 1967	408,250 ± 120
	I.A.U. (1966)	408,000
	J.P.L. (1969)	408,522 ± 3
Mars	Clemence	3,088,000 ± 300
	Rabe 1954	3,110,800
	Ash, Shapiro and Smith 1967	3,111,000 ± 9000
	Null Gordon Tito 1967	3,098,600 ± 600
	I.A.U. (1966)	3,903,500
	J.P.L. (1969)	3,098,700 ± 100
Jupiter	Clemence 1961	1047.39 ± 0.03
	O'Handley 1967	1047.387 ± 0.008
	I.A.U. (1966)	1047.355
	J.P.L. (1969)	1047.3908 ± 0.0074
Saturn	Hertz 1953	3497.64
	van der Bosch 1927	3496.0 ± 3
	Clemence 1960	3499.7 ± 0.4
	I.A.U. (1966)	3501.6
	J.P.L. (1969)	3499.2 ± 0.4
Uranus	Hill 1898	22,239 ± 89
	Harris 1950	22,934 ± 6
	I.A.U. (1966)	22,869
	J.P.L. (1969)	22,930 ± 6
Neptune	Gaillot 1910	19,094 ± 22
	van Biesbroeck 1957	18,889 ± 62
	I.A U. (1966)	19,314
	J.P.L. (1969)	19,260 ± 100
Pluto	Brouwer 1955	400,000 ± 40,000
	I.A.U. (1966)	360,000
	J.P.L. (1969)	1,812,000 ± 40,000

The IAU system of constants is taken from the January 1966 Supplement to the Astronomical Ephemeris[26]. The table on the next page gives the NASA adopted values as discussed by Clarke[4] in 1964. Mulholland[25] has recently taken a list of the best available measurements, (which includes most of the above list excepting those of the IAU and Rabe) and recommends a set of ratios which represents the weighted mean of these data. The values are weighted according to their own internal standard deviations. This is certainly a logical approach but these standard deviations represent the internal consistency of the measurement and are only incidentally related to accuracy. However, the only other weighting process one can use is personal judgment.

Some additional planet constants are given below. They are taken from Lockheed Rpt LR17571 (AFCRL-63-705) by Mary P. Francis[8].

Planet	Equatorial Radius (km)	J_2	J_4	JPL - 1969 Equatorial Radius
Mercury	2424 ± 11	0	0	2435.0 ± 3.0
Venus	6100 ± 18	0	0	6052.0 ± 3.0
Mars	3412 ± 4	.001916	0 ± 0.0001	3393.4 ± 4.0
Jupiter	71420 ± 50	.00735	-.000675	71372.0
Saturn	60440 ± 250	.01667	-.001 ± .0005	60401.0
Uranus	24860 ± 200	-	-	23535.0
Neptune	26500 ± 2000	.005	0 ± .001	22324.0
Pluto	4000 ± 2000	-	-	7016.0

Table of Planetary Mass Ratios

Planet	NASA 1964[4]	Mulholland 1968[25]	J.P.L. 1969[30]
Mercury	6,120,000	6,017,000	5,983,000
Venus	408,539.5	408,504	408,522
Earth-Moon	328,905.8	328,900	328,900.1
Mars	3,088,000	3,098,500	3,098,700
Jupiter	1047.39	1047.3908	1047.3908
Saturn	3500	3499.2	3499.2
Uranus	22,869	22,930	22,930
Neptune	18,889	19,071	19,260
Pluto	400,000	400,000	1,812,000

Except for Pluto, the order of the differences are a pretty good indication of the uncertainties. The new value of Pluto was determined by R.L.Duncombe, W.J.Klepczynski and P.K.Seidelmann of the Naval Observatory in August 1968.

Reference 31 gives a rather complete list of physical constants and conversion factors. Since 1 July 1959 the following standards apply:

1 foot = 0.3048 meters (exactly)
1 meter = 3.280839895 feet
1 nautical mile = 1852 meters (exactly)
1 nautical mile = 6076.115486 feet

The previous foot standard is now called the U.S. Survey foot and is 1200/3937 meters (exactly). One nautical mile is thus 6076.103333 U.S. Survey feet.

Some other constants of interest are given below

Planet	Sidereal Period of Rotation[30]	Orbital Period	(km) Uncertainty Estimates for Planetary Position[30]
Mercury	58.67 ± 0.03d	87.969d	200 - 300
Venus	242.6d	224.701d	300 - 500
Earth-Moon	$23^h56^m4^s.1$	365.256d	150 - 300
Mars	$24^h37^m22^s.6689$	1.88089 yr	200 - 400
Jupiter	$9^h50^m30^s.003$	11.86223 yr	2000
Saturn	10^h14^m	29.45774 yr	2000
Uranus	10^h49^m	84.018 yr	3000
Neptune	14^h	164.78 yr	6000
Pluto	6.39d	248.4 yr	2×10^6

REFERENCES

1. Anderson, J.D., C. W. Null and C. T. Thornton, "Evaluation of Astro Constants from Tracking of Mariner II," Vol 14 of Progress in Astronautics and Aeronautics, Academic Press, 1964.

2. D. Brouwer and G. M. Clemence, Chapter 3, "Orbits and Masses of Planets and Satellites," p 31 of the book "Planets and Satellites" edited by C. P. Kuiper and B. M. Middlehurst.

3. D. Brouwer, "System of Astronomical Constants," p 115, Vol 14, Progress in Astronautics and Aeronautics, "Celestial Mechanics and Astrodynamics," Academic Press 1964.

4. V. C. Clarke, Jr., "Constants and Related Data for Use in Trajectory Calculations, as Adopted by the Ad Hoc NASA Standard Constants Committee," Tech Rpt 32-604, Jet Propulsion Lab, Pasadena, Calif., 6 March 1964.

5. de Sitter, W. and D. Brouwer, "On the System of Astronautical Constants," Bulletin of Astronomical Soc. Netherlands (B.A.N.), 8 July 1938, p 213-231.

6. K. D. Froome, "Precision Determination of the Velocity of Electromagnetic Waves," Nature, Vol 181, No. 4604, p 258 (1958).

7. Irene Fischer, "An Astrogeodetic World Datum from Goidal Heights Based on Flattening of $\frac{1}{298.3}$," Journal of Geophysical Research, Vol 65, No. 2067, 1960.

8. Francis, M. P., "Astrodynamic Constants Analysis," Lockheed California Comp. AFCRL-63-705, July 1963.

9. Herrick, S., R. M. L. Baker, Jr. and C. C. Hilton, "Gravitational and Related Constants for Accurate Space Navigation," Proceeds of 8th International Astronautical Congress Barcelona 1957.

10. Jones, H. Spencer, "The Solar Parallax and the Mass of the Moon from Observations of Eros at Opposition of 1931," Mem. Royal. Astrom. Soc., Vol 66, Part II, 1941.

11. Kotelnikov, V. A., "Radar Contact with Venus," Paper presented at 12th International Astronautical Congress, Washington, D.C., Oct 1961.

12. Makemoon, M. W., R. M. L. Baker, Jr. and C. B. Westrom, "Analysis and Standardization of Astrodynamic Constants," Journal of Astronautical Science, Vol VIII, No. 1, Spring 1961.

13. Marsden, B. C., "An Attempt to Reconcile the Dynamical and Radar Determination of the A.U.," I.A.U. Symposium No. 21 on Fundamental Constants of Astronomy, Paris, June 1963.

14. Muhleman, D. O., D. B. Holdridge, N. Block, "The Astronomical Unit Determined by Radar Reflections from Venus," Astrom. Journal, Vol 67, May 1962, p 191-200.

15. Pettengill, E. H., et. al., "A Radar Investigation of Venus," Astrom. Journal, Vol 67, May 1962, p 181-190.

16. Rabe, E., Derivation of the Fundamental Astronomical Constants from Eros during 1926-1945. Astrom. Journal, Vol 55 (1950), p 112 and Vol 59 (1954), p 409-411.

17. Jones, Spencer, Chapter 2 of Vol II, "The Earth as a Planet," edited by G. P. Kuiper, University of Chicago Press.

18. Thompson, J. H., et. al., Nature, Vol 190 (1961), p 519.

19. Wolaver, L. E., Space Technology Lecture Series, WADC Report TR-59-732, Vol I, Oct 1958, p 77-99, "Factors in Precision Space Flight Trajectories.

20. Mercer, R. J., "Physical Constants for Satellite Calculations," TDR 469(5110-02)-2, Jan. 1965, Aerospace Corp.

21. Yaplee, B. S., H. Knowles, H. Shapiro, D. J. Craig and D. Brouwer, "The Mean Distance to the Moon as Determined by Radar," Paper at Symposium No. 21 on Fundamental Constants of Astronomy, Paris, June 1963.

22. Ash, M. E., Shapiro, I. I. and Smith, W. B., "Astronomical Constants and Planetary Ephemerides Deduced from Radar and Optical Observations" Astronomical Journal vol 22 pp 338-350, 1967.

23. Sjogren, W. L., Trask, D. W., Vegos, C. J. and Wollenhaupt, W. R., "Physical Constants as Determined from Radio Tracking of the Ranger Probes" Technical Report 32-1057, JPL, Pasadena, California December 30, 1966.

24. Null, G., Gordon, H. J. and Tito, D. A., "Mariner IV Flight Path and Its Determination from Tracking Data" Technical Report 32-1108, JPL 1968.

25. Mulholland, J. D., "Planetary Masses for Ephemeris Development" JPL Space Programs Summary 37-45, vol 4, pp 17-19, 1968.

26. Anon, "The IAU System of Astronomical Constants" Supplement to the Astronomical Ephemeris Prepared jointly by U.S. Naval Observatory Nautical Almanae Office and the Royal Greenwich Observatory, January 1966 (AD-660979).

27. Bray. T. A., Goudas, C. L., and Zdenka Kopal, "Estimates of the zonal gravity harmonics of the moon." Incarus, Vol 7, No. 1 pp 76-84, July 1967.

28. Francis, Mary F. and Eugene Rabe, "The Earth + Moon Mass and other Astronomical Constants from the Eros Motion 1926 - 1965." Astronomical Journal, Vol 72, No. 7, Sept. 1967, p 856 - 864.

29. Rabe, Eugene, "Corrected Deviation of Astronomical Constants from the Observations of Eros 1926 - 1945," Astronomical Journal, Vol 72, No. 7, Sept 1967, p 852 - 855.

30. Melbourne, W. G.; Mulholland, J. P.; Sjozren, W. L.; Sturms, F. M.; "Constants and Related Information for Astrodynamic Calculations, 1968." J.P.L. Technical Report 32 - 1306, July 15, 1968.

31. Mechlty, E. A., "The International System of Units - Physical Constants and Conversion Factors." NASA . SP-7012 (1964).

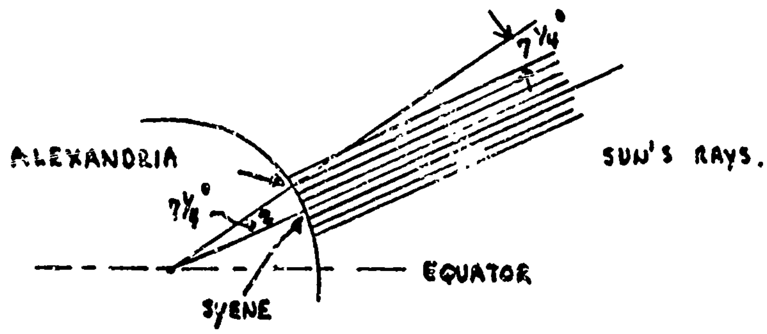
35. THE SHAPE OF THE EARTH

Let's start with the flat earth. How do we know - how did the ancients know - that the earth is round? Three phenomena alert us to the earth's shape. (1) During eclipses of the moon; the sun, earth and moon are in a straight line and hence the circular shadow which creeps over the moon is the earth's shadow. (2) The sun appears to climb higher in the sky as one travels further south and new stars become visible, thus indicating that the earth's surface is curved. (3) The way in which ships disappear over the horizon. The Flat Earth Society, a still active group whose president has gone around the world lecturing on the flat earth, used to write letters to Life magazine whenever they printed pictures showing a curved horizon - obviously faked photographs! So it is no wonder the ancients were slow to accept a spherical earth.

The credit for first proposing the idea that the earth is spherical is usually given to Pythagoras in sixth century B.C.[11] He seems to have gotten the idea from his conviction that the sphere was the perfect shape. The basis for the belief was shakey but none the less the idea was right and was adopted by later influential Greek philosophers including Plato and the archvillian Aristotle.

Eratosthenes, in the 3rd century B.C., did more than adopt the idea, he made the first measurement of the earth's circumference. He noticed that at noon at midsummer the sun was directly overhead at Syene, or

Aswan as it is now known, while at Alexandria almost due north, it was $\frac{1}{50}$ of a circle (7.25°) away from the vertical.



So between Alexandria and Syene the direction of the earth's surface changes by about 1/50 of a circle. To turn through a complete circle, around the world, one must travel 50 times the distance between Alexandria and Syene. Eratosthenes did not really know what this distance was, but, being a resourceful man, he noted that it took 50 days to do the journey by camel and that camels usually travelled about 100 stadia per day. Hence the distance between Alexandria and Syene was 5000 stadia, so the circumference of the earth was 250,000 stadia. This corresponds either to 46,000 km or 39,700 km depending on the definition of the stadium. The correct value is just over 40,000 km so Eratosthenes did very well indeed. Poseidonius using the star Canopus and the distance from Rhodus to Alexandria improved the value. He obtained the distance down the Nile from Rhodus to Alexandria by using the sailing time instead of camel time. Caliph Abdullah at Mamum used wooden rods to measure the set distance in 800 A.D. to arrive at a figure only 3.6% too great. This method in refined form, is still one of the best methods of measuring

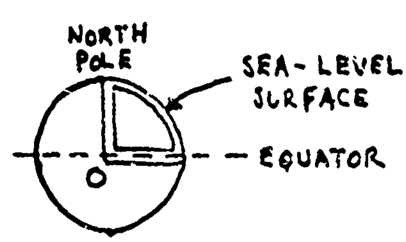
the earth's radius. Picard in 1671 measured a long arc in Northern France and gave us the first accurate value.

Actually the earth is better approximated by an oblate spheroid, a sphere which is flattened at the poles and has an elliptic cross section. The flattening ratio f is defined by

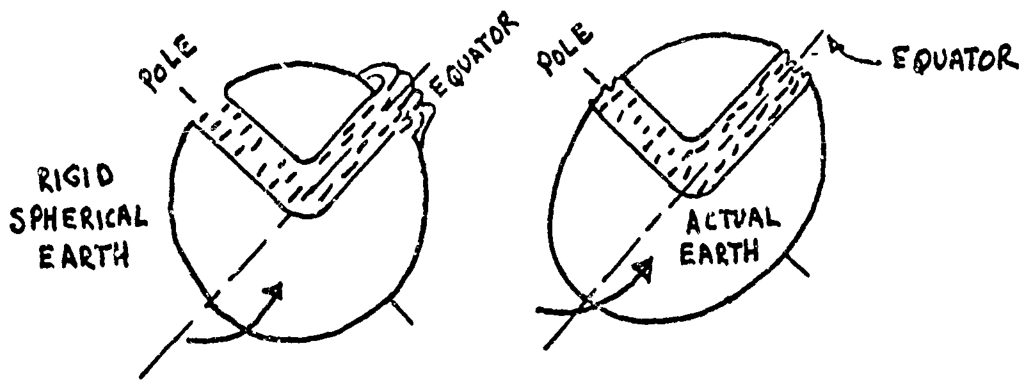
$$f = \frac{\text{Equatorial Diameter Minus Polar Diameter}}{\text{Equatorial Diameter}} \qquad (35\text{-}1)$$

The first indication that the earth might be flattened at the poles was obtained in 1672 by a French expedition to South America, led by Jean Richer.[11] He noticed that a pendulum clock which kept good time at Paris lost 2 1/2 minutes per day at Cayenne. It was suspected that these measurements indicated the earth was not exactly spherical, but no numerical values could be obtained because Newton's law of gravitation had not then been published.

The first numerical estimate of the earth's flattening was made by Newton from purely theoretical arguments (Principia 1687[5]). Newton imagine one tube of water, or canal as he called it, running from the North Pole to the earth's center, and another from the center to a point on the equator.



Since water shows no tendency to run from equator to pole (or vice versa) over the surface, the pressure exerted by each canal at the earth's center must be equal. Because of the earth's rotation, gravity is slightly less in the equatorial canal, which will therefore be longer. Proceeding this way, Newton estimated the flattening as $\frac{1}{230}$. The value is a little high because he made no allowance for any increase in density toward the center of the earth.



On spherical earth, water would spill out. Although no canal exists, the oceans correspondingly would flow to the equator. Since they do not - q.e.d.

The interior of the earth is more complicated. The mean density of the earth is 5.52 but the superficial layer has a density of 2.71. Beneath this veneer of sedimentary rock is a granite layer usually 10 to 20 km thick under continents and thinner under the Antlantic and Indian Oceans - probably absent under most of the Pacific. Under this is a denser basalt layer about as thick. These layers form the crust and are but 1% of the earth's mass. Below the crust at a depth of

30 to 50 km from the surface is a fairly abrupt transition to the mantle where the density is from 3.5 to 5.5 and is probably composed of basic silicate rock. The jump in density is called the Mohorovicic discontinuity. The mantle extends to about 2900 km and accounts for about 70% of the earth's mass. The central core some 3400 km in diameter is even denser, 10 to 12 gm/cm^3 and is believed composed of metallic iron alloyed with nickel.

The pressure here is very great - estimated to be 1 1/2 to 3 x 10^6 atmospheres. This interior is more rigid than steel and has a very high temperature.[28]

The problem of the elliptical shape of the earth was settled by making long accurate arc measurements at various parts of the globe. One of the first precise measurements was made by Picard in Northern France in 1671, the same measurement which relaunched Newton on his gravitational theory. Today there are vast networks of triangulation nets which connect surveys. The figure on the next page shows the spacing and distribution of astro-geodetic geoid data. Some areas are very well mapped and long arcs have been accurately measured but the data is relatively sparse on a world wide basis.

In order to represent the figure of the earth, or more specifically the shape of its equipotential surface, the geoid, one assumes an ellipsoid of revolution and measures small deviations from this main figure. The particular ellipsoid which is used by any one country is called the reference ellipsoid and unfortunately they do not all use the same ellipsoid. The following are ellipsoids in current use:

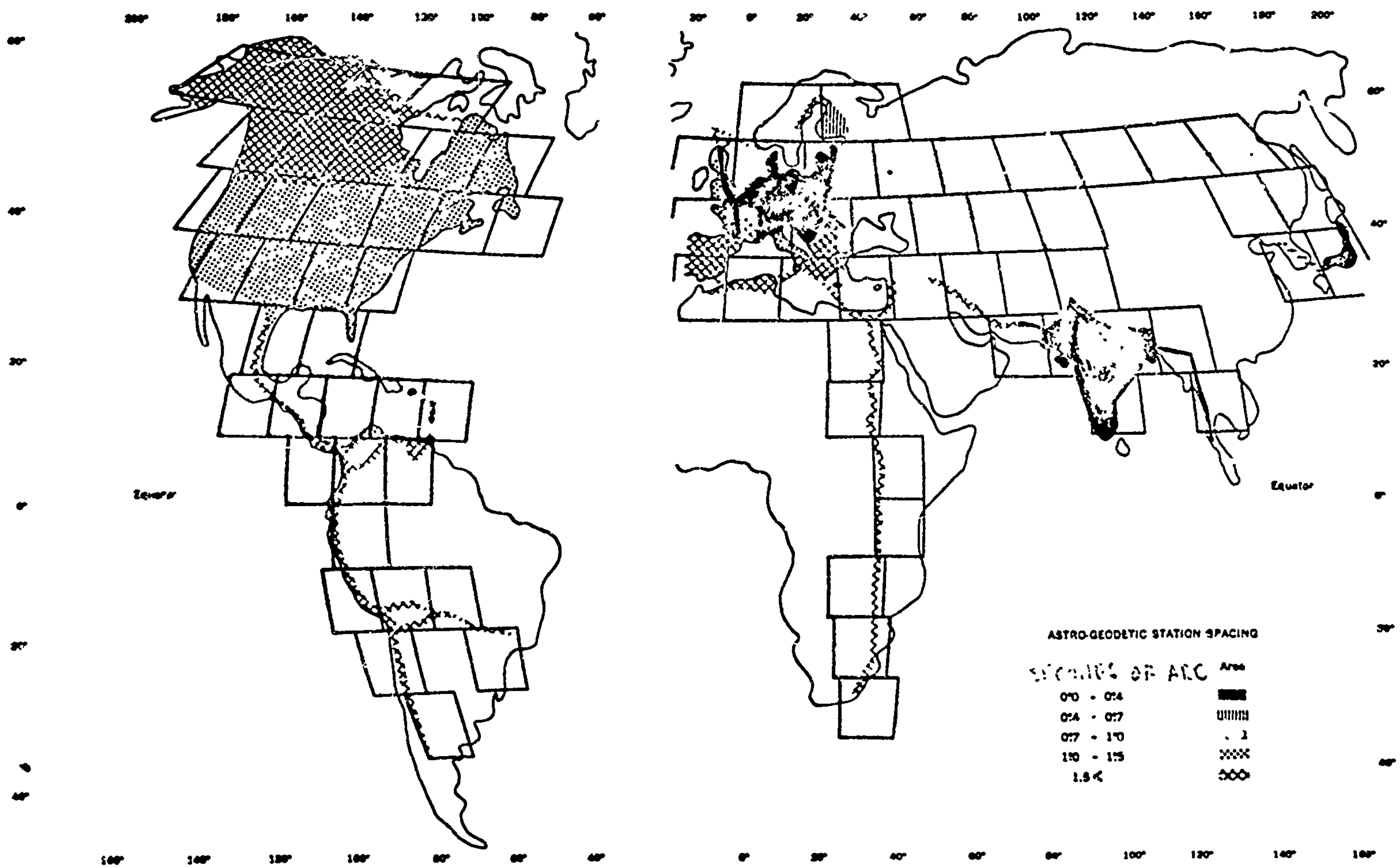


Figure 1 - Spacing and distribution of astro-geodetic geoid data

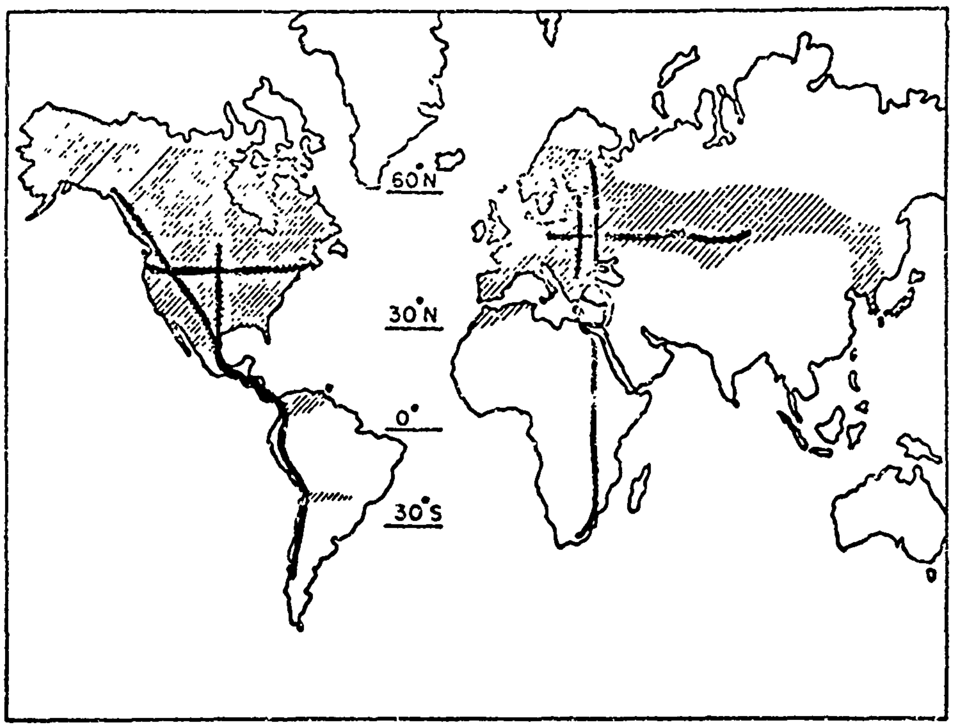


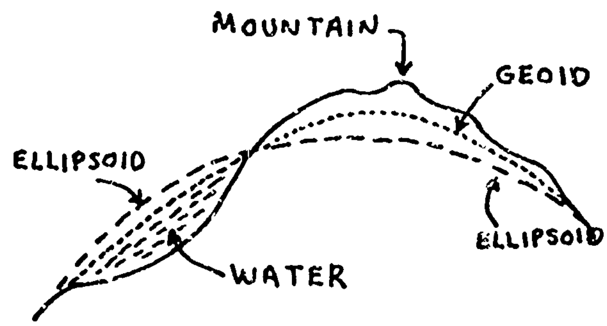
FIG. 1A Arcs used in new determination of the figure of the earth

The International Ellipsoid was developed by Hayford in 1910 and adopted by the International Union of Geodesy and Geophysics. In the list below the names of the individuals who derived the ellipsoids and year of development is given.

NAME	a_e	$\frac{1}{f}$	WHERE USED
Krassowksy 1940	6378245	298.3	Russia
International 1924	6378388	297.0	Europe
Clarke 1880	6378249	293.5	France, Africa
Clarke 1866	6378206	295.0	North America
Bessel 1941	6377397	299.15	Japan
Airy 1830	6376542	299	Great Britain
Everest 1830	6377276	300	India
Hough 1956	6378260	297.0	Army Map Service
Kaula 1961	6378163	298.24	NASA

The difference in dimensions are due principally to different arc lengths used as a reference. These differ at points on the earth because the shape of the geoid varies. The figure on the next page indicates how the mass surplus of the mountains and the mass deficiencies of the ocean cause the geoid to differ from any reference ellipsoid.

An equipotential surface is one which is everywhere perpendicular to a pull or an acceleration: the geoid is that equipotential surface which most nearly coincides with the mean sea level of the earth.



In addition there are other complications. When we are near the mountains, the excess mass deflects the plumb-bob vertical as compared to the astronomical vertical, as is expected. However, as we move away from the mass, the amount of the mass deflection of the vertical falls off much faster than an application of the simple inverse-square law of attraction would indicate. It now appears that the density of mass below the mountain is correspondingly less, thereby compensating for the excess mass of the mountain above the surface. As we move away from the mountain, the excess mass effect vanishes. Similarly there is an increase in density in the mass below the waters. These must be taken into account in determining the geoid and the technique is referred to as isostatic reduction. This isostasy theory holds good on the average, but there are exceptions (notably the Island of Cyprus and parts of the ocean near East and West Indies). Therefore, we see that not only does an ellipsoid of revolution or any geometrical figure fail to represent the figure of the geoid, but we are not too sure where the geoid is exactly located.

One way to represent this potential field of the earth is to consider it a sphere with "density-bumps". Just as we decompose data into its

Fourier Series components, we can develop the geoid by representing it in spherical harmonics of some sort. They must be harmonic functions because the potential function satisfies La Place's equation. The polynomials used are called Legendre Polynomials;[12] they arise quite naturally in developing the potential function.

$$U = \frac{k^2 M_e}{r}\left[1 - \sum_{n=2}^{\infty} J_n \left(\frac{a_e}{r}\right)^n P_n \,(\sin L)\right] \qquad (35\text{-}2)$$

where k^2 is Gaus's constant, M_e, the mass of the earth so $\mu = k^2 M_e = 3986032\ \mathrm{km^3/sec^2}$; a_e is the earth's equatorial radius = 6378166 km, the J_n are constants to be determined, P_n are Legendre Polynomials of degree n and L is the instantaneous latitude of the satellite. The J_1 term is zero if the equatorial plane is chosen to pass through the earth's center of mass. The explicit form of some of these Legendre Polynomials are listed as follows:

$$P_2 = \frac{1}{2}\,(3 \sin^2 L - 1)$$

$$P_3 = \frac{1}{2}\,(5 \sin^3 L - 3 \sin L)$$

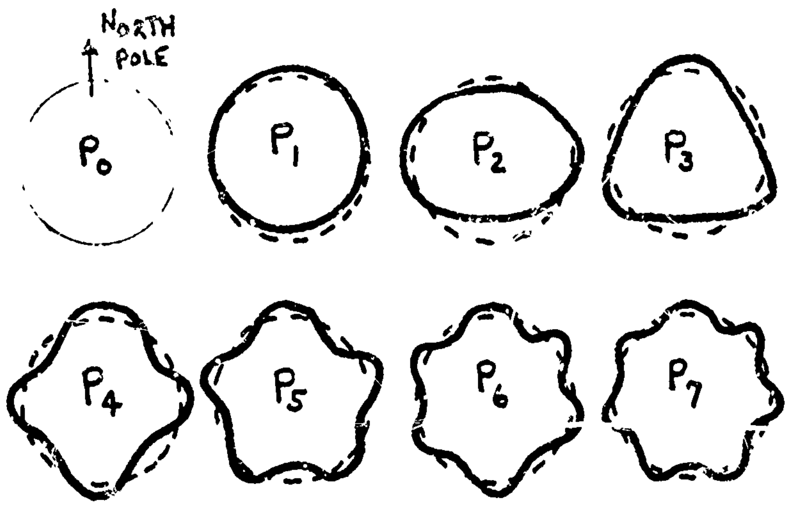
$$P_4 = \frac{1}{8}\,(35 \sin^4 L - 30 \sin^2 L + 3)$$

$$P_5 = \frac{1}{8}\,(63 \sin^5 L - 30 \sin^3 L + 15 \sin L)$$

$P_6 = \frac{1}{16} [231 \sin^6 L - 315 \sin^4 L + 105 \sin^2 L - 5)$

etc.

The largest term is the coefficient J_2 which results from the earth's flattening. The next term, J_3, corresponds to a tendency toward a triangular shape (see below) or a pear shape as it is often called because if the tendency were stronger the curve would become concave like a pear. The next term might be called square-shaped, the fifth harmonic gives rise to five "petals," and so forth -



(ROUGH SKETCH)

etc.

By superimposing enough of these harmonics it is possible to build up any shape and any gravitational field (with the limitation that they must be symmetrical about the polar axis).

We hope higher terms can be safely neglected. In regard to the potential function as given in equation (2), some earlier representations[3] which used only the first three terms used the notation

$$J = \frac{3}{2} J_2 \qquad H = \frac{5}{2} J_3 \qquad D = -\frac{35}{8} J_4$$

and still others [30] used $K_2 = -\frac{1}{2} J_2$, $K_4 = \frac{J_4}{8}$ and other variations abound. The J_n notation is now widely accepted.

The best way to determine these J_n values is by satellite motion. The equations of motion using the potential of equation (2) may be written down, expanded and integrated term by term. A process much easier to say than do. We can expand the perturbation to the central force field in powers of J_n and functions F of the orbital elements as

$$(4) \qquad P = \sum_{n=2}^{\infty} J_n F_n(a,e,i,\omega) + \sum_{n=2}^{\infty} \sum_{m=2}^{\infty} J_n J_m F_{nm}(a,e,i,\omega) + \cdots$$

Since the J_n are of order 10^{-6} for $n>2$ it has only been necessary to evaluate the terms in the first series, together with only the J_2^2 terms in the second series.

This orbital theory has been developed by various authors including King-Hele,[33] Kozai,[35] Brouwer,[30] Groves,[32] Merson,[36] etc., using

different methods and they all seem to agree. The most generalized version is that of Groves[32] [1960] and the most complete explicit forms are those of Merson[36] (1960).

Merson's results, with a few negligible terms dropped and with the addition of J_2^2 terms from King-Hele,[9] are given below: $[f = \sin^2 i]$

$$\frac{d\bar{\Omega}}{dt} = -\frac{\sqrt{\mu}}{a^{3/2}}\left(\frac{a_e}{p}\right)^2 \cos i \left[\frac{3}{2} J_2 - \frac{3}{2} J_3 \frac{a_e}{p} \frac{e \sin \omega}{\sin i}\left(1 - \frac{15}{4} f\right)\right.$$

$$-\frac{15}{4} J_4 \left(\frac{a_e}{p}\right)^2 \left\{\left(1 - \frac{7}{4} f\right)\left(1 + \frac{3}{2} e^2\right) + \left(\frac{7}{4} f - \frac{3}{4}\right) e^2 \cos 2\omega\right\}$$

$$+\frac{15}{4} J_5 \left(\frac{a_e}{p}\right)^3 e \operatorname{cosec} i \left\{\left(1 - \frac{21}{2} f + \frac{105}{8} f^2\right)\left(1 + \frac{3}{4} e^2\right) \sin \omega\right.$$

(35-5)

$$\left. + \frac{7}{8}\left(1 - \frac{15}{8} f\right) e^2 \sin 3\omega\right\} + \frac{105}{16} J_6 \left(\frac{a_e}{p}\right)^4 \left\{\right.$$

$$\left(1 - \frac{9}{2} f + \frac{33}{8} f^2\right)\left(1 + 5e^2 + \frac{15}{8} e^4\right) - \frac{5}{2}\left(1 - 6f + \frac{49}{16} f^2\right)\left(1 + \frac{1}{2} e^2\right) e^2 \cos 2\omega$$

(continued on next page.)

$$- \frac{15}{32}\left(1 - \frac{33}{20} f\right) f e^4 \cos 4\omega \Bigg\}$$

$$- \frac{315}{32} J_8\left(\frac{a_e}{p}\right)^6\left(1 - \frac{33}{4} f + \frac{143}{8} f^2 - \frac{715}{64} f^3\right) + 0\left(J_7, J_8 e^2, J_9 \cdots\right)$$

$$- \frac{9}{4} J_2^2 \left(\frac{a_e}{p}\right)^2 \left(1 - \frac{19}{12} f\right) + 0\left(J_2^2 e^2, J_2 J_3, J_2 J_4, \cdots\right)\Bigg] .$$

$$\frac{d\bar{\omega}}{dt} = \frac{\sqrt{\mu}}{a^{3/2}}\left(\frac{a_e}{p}\right)^2 \left[3J_2 \left(1 - \frac{5f}{4}\right) + \right.$$

(35-6)

$$\left\{1 - \frac{5f}{4} + e^2 \left(\frac{35}{4} - \frac{35f}{4} - \frac{1}{f}\right)\right\} \frac{a_e}{p} \frac{3}{2} J_3 \frac{\sin i \sin\omega}{e}$$

$$- \frac{15}{2} J_4 \left(\frac{a_e}{p}\right)^2 \left\{1 - \frac{31}{8} f + \frac{49}{16} f^2 + f \cos 2\omega \left(\frac{3}{8} - \frac{7}{16} f\right)\right\}$$

$$- \frac{15}{4} J_5 \left(\frac{a_e}{p}\right)^3 \frac{\sin i \sin \omega}{e} \left(1 - \frac{7f}{2} + \frac{21}{8} f^2\right)$$

(continued on next page.)

$$+ \frac{105}{8} J_6\left(\frac{a_e}{p}\right)^4 \left\{\left(1 - 8f + \frac{129f^2}{8} - \frac{297f^3}{32}\right)\right.$$

$$\left. + \left(\frac{5}{4} - \frac{15}{4} f + \frac{165f^2}{64}\right) f \cos 2\omega\right\} -$$

$$- \frac{315}{4} J_8 \left(\frac{a_e}{p}\right)^6 \left(1 - \frac{331}{32} f + \frac{2057}{64} f^2 - \frac{19877}{512} f^3 + \frac{16445}{1024} f^4\right)$$

$$+ O(J_4 e^2, J_5 e^2, J_6 e^2, J_7, J_8 \cos 2\omega, J_8 e^2, J_9, J_{10}, \cdots)$$

$$+ O(J_2{}^2, J_2 J_3, \cdots)$$

where $p = a(1-e^2)$, $f = \sin^2 i$. The value of a is found from the nodal period T_N by the equation

$$T_N = 2\pi \sqrt{\frac{a^3}{u}} \left\{1 - \frac{3J_2}{8} \left(\frac{a_e}{p}\right)^2 (7 \cos^2 i - 1) + O(J_2 e)\right\} .$$

In these equations the orbital elements, a, e and i are not taken as osculating elements but as certain mean elements. The need for an exact definition of the elements only arises in the J_2^2 term, whose form is completely different if, for example, osculating elements evaluated at the ascending node are used.

Note that some of the terms in (5) and (6) are not strictly "secular," since they contain the periodic factors cos ω, cos 2ω, etc. It is desirable to include them in the same formulae because ω may change only a small fraction of a cycle during the lifetime of a satellite, thus allowing the long-period terms to masquerade successfully as secular terms.

Since a is constant, the oscillations in e can most easily be visualized as a change in the perigee distance from the earth's center $r_p = a(1-e)$. From Merson we have

$$i = i_o + \frac{eJ_3a_e}{2J_2p} \cos i \sin \omega + \frac{15J_4e^2}{64J_2}\left(\frac{a_e}{p}\right)^2 \frac{1 - \frac{7}{6} f}{1 - \frac{5}{4} f} \sin 2i\left(1-\cos 2\omega\right)$$

$$- \frac{5eJ_5}{4J_2}\left(\frac{a_e}{p}\right)^2 \left(1 - \frac{7}{2} f + \frac{21}{8} f^2\right)\left(1 - \frac{5}{4} f\right) \cos i \sin \omega$$

(35-7)

$$+ 0 \left(J_6e^2, J_2^2e\right)$$

$$r_p = r_{p_o} + \frac{J_3 a_e}{2J_2} \sin i \sin \omega + \frac{15 e J_4 a_e f}{32 J_2} \frac{a_e}{p} \frac{1 - \frac{7f}{6}}{1 - \frac{5f}{4}} (1 - \cos 2\omega)$$

$$- \frac{5J_5 a_e}{4J_2} \left(\frac{a_e}{p}\right)^2 \left(1 - \frac{7}{2} f + \frac{21}{8} f^2\right) \left(1 - \frac{5}{4} f\right)^{-1} \sin i \sin \omega$$

(35-8)

$$+ 0\left(J_5 e, J_2^2\right).$$

where suffix 0 denotes values when $\omega = 0$.

For an orbit with a rapidly rotating major axis, the periodic terms in equations (5) to (8) yield equations for J_3 and J_5. The contributions from J_4 can be excluded, since they have a different period and moreover are small as a result of the factor e. For example equation (8) shows that the distance of perigee from the earth's center oscillates with the same period as ω and amplitude

$$\frac{a_e}{2} \sin i \left\{\frac{J_3}{J_2} - \frac{5}{2} \frac{J_5}{J_2} \left(\frac{a_e}{p}\right)^2 \left(1 - \frac{7}{2} f + \frac{21}{8} f^2\right) \left(1 - \frac{5f}{4}\right)^{-1}\right\}$$

Thus the observed amplitude of oscillation of each orbital element yields an equation for J_3, J_5, ··· from which the values of the odd coefficients can be found.

The strictly secular terms in equations (5) and (6) can be used to evaluate the even J_n. The contributions of the odd J_n to equations (5) and (6) are small, because of multiplying factors of the form $e \sin \omega$, etc., which are of order 0.1. If the odd J_n terms are still appreciable, their influence can be further reduced by choosing a time interval during which the mean value of $\sin \omega$ is small. Consequently, the J_3 and J_5 term in equation (5) and (6) can often be ignored; and if not, the imperfectly-known values of J_3 and J_5 can be substituted without introducing appreciable errors.

Thus equation (5) with an observed value of $\frac{d\Omega}{dt}$ and an approximate value for J_2^2 provides in effect one linear relation between the even J_n.

If there are r observed values of $\frac{d\bar{\Omega}}{dt}$ available, we have r equations between the J_n from which the values J_2, J_4, $\cdot\cdot$ J_{2r} can in principle be determined if J_{2r+2}, J_{2r+4}, $\cdot\cdot$ are assumed zero. In practice, the observed values must be derived from satellites with appreciably different values of either i or p; otherwise the simultaneous equations will tend to be ill-conditioned.

Similar conclusions apply to equation 6. If however, a total of r values of $\frac{d\bar{\Omega}}{dt}$ and $\frac{d\bar{\omega}}{dt}$ are available from suitably chosen satellites, it should be possible to evaluate J_2, J_4, $\cdot\cdot$ J_{2r}.

To sum up, equations (7) and (8) and the periodic terms in (5) and (6) serve to determine the odd J_n, while the secular terms in (5) and (6) give the even J_n. When more accurate orbits are available in large number

it should be possible to evaluate a large number of the J_n. We shall comment on this shortly.

The observed values for use in these equations must, of course, be purged of any contributions from other sources of perturbations. G. E. Cook[31] has developed an expression for $\frac{d\bar{\Omega}}{dt}$ and $\frac{d\bar{\omega}}{dt}$ due to an oblate, rotating atmosphere in which the air density varies exponentially with height. Kozai[34] has given these same effects due to the moon and sun's attraction. These latter formulae are used to correct the observed variations in Ω and ω. The other perturbations due to relatively, electromagnetic forces, effect of planets and solar radiation pressure[38] are negligible to the order of accuracy so far achieved (Echo satellites excluded).

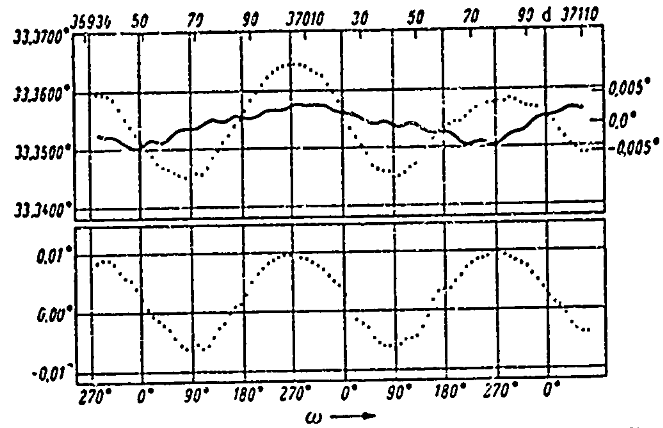
To illustrate the range of these variables we reproduce some results of Kozai based on precisely reduced observations of Satellite 1959 Eta (Vanguard III) during 178 days from Jan 1 to June 27 of 1960.

In the upper halves of Figures 2 through 5 the observed values of four orbital elements are plotted every two days. On Figures 4 and 5 the variations due to the secular motions have been subtracted from the observed values by

$$\Omega - (106^\circ\!.1189 - 3.273988t)$$

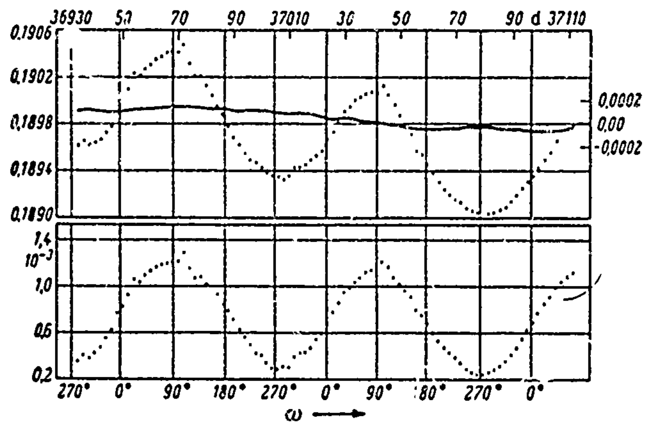
$$\omega - (138^\circ\!.4484 + 4.874325t)$$

The lunisolar gravitational perturbations of long period were computed by Kozai[34] and are shown by solid lines in the figures. The



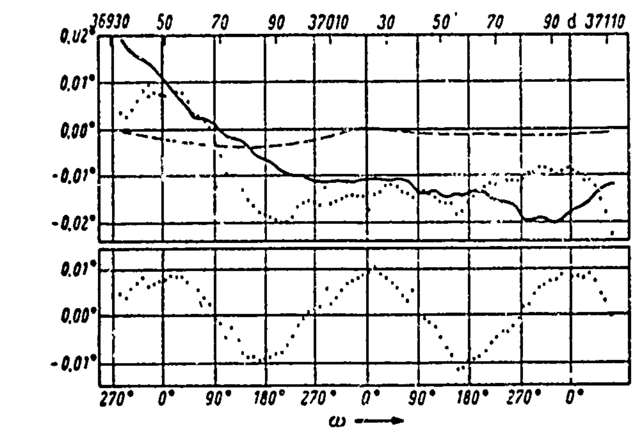
2

Variations of inclination. Lunisolar perturbation is expressed by a solid line with a scale on the right side.



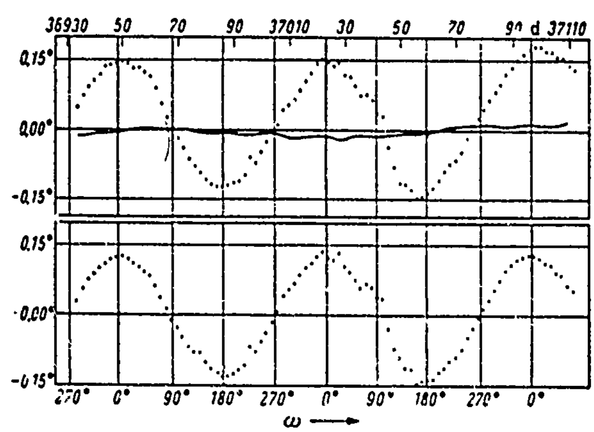
3

Variations of eccentricity. Lunisolar pertubation is expressed by a solid line.



4

. Variations of longitude of ascending node. Lunisolar perturbation is expressed by a solid line and solar radiation effect by a broken line.



5

Variations of argument of perigee. Lunisolar perturbation is expressed by a solid line.

solar radiation pressure effects based on $\frac{A}{M} = 0.173$ (cgs) are too small to be plotted except for the longitude of the ascending node (Ω) in Figure 4 where it is expressed by a broken line.

After subtracting the air-drag, lunisolar, solar-radiation pressure perturbations the long periodic terms due to even harmonics in the potential are calculated using approximate values of the even harmonics. These latter corrections are given by

$$\delta e = -0.23 \times 10^{-4} \cos 2\omega$$

$$\delta i = 0\overset{\circ}{.}40 \times 10^{-3} \cos 2\omega$$

$$\delta \omega = 0\overset{\circ}{.}67 \times 10^{-2} \sin 2\omega$$

$$\delta \Omega = 0\overset{\circ}{.}75 \times 10^{-3} \sin 2\omega$$

After subtracting all of these perturbations the corrected values of the orbital elements are plotted in the lower halves of Figures 2 through 5. The corrected variations may be regarded as those due to the odd-order harmonics in the potential. Note that the observed values of i and Ω are significantly distorted by lunisolar perturbations.

By fitting the corrected observed values by $\sin \omega$ and $\cos \omega$ plus a secular term, Kozai derives the following set of elements:

Epoch Feb 14.0, 1960

$n = 11.06540 \pm 1$

$i = 33.35428 \pm 6 - (0\overset{\circ}{.}763 \pm 9) \times 10^{-2} \sin \omega$

$e = 0.189755 \pm 3 + (0.459 \pm 5) \times 10^{-3} \sin \omega$

$\Omega = 106\overset{\circ}{.}1189 \pm 2 - (3.273988 \pm 7)\, t + (0\overset{\circ}{.}89 \pm 2) \times 10^{-2} \cos \omega$

$\omega = 138\overset{\circ}{.}4484 \pm 10 + (4\overset{\circ}{.}87433 \pm 4)\, t + (0\overset{\circ}{.}1267 \pm 12) \cos \omega$

We seem to have six significant figures for the mean element except for ω and e which have 5.

Kozai used $a_e = 6.378165 \times 10^8$ cm and $k^2 M_e = 3.986032 \times 10^{20}\ \mathrm{cm}^3/\mathrm{sec}^2$. By including spherical harmonics up to ninth order, the nominal (mean) values of the orbital elements can be used to write the equations of conditions as follows:

$e = -\,190.5J_3 + 84.2J_5 + 50.7J_7 - 71.4J_9 = (0.459 \pm 5) \times 10^{-3}$

$i = 3264J_3 - 1443J_5 - 869J_7 + 1223J_9 = -\,(0\overset{\circ}{.}763 \pm 9) \times 10^{-2}$

$\omega = -\,54711J_3 + 41712J_5 + 10806J_7 - 39020J_9 = 0\overset{\circ}{.}1267 \pm 12$

$$\Omega = -5937J_3 - 11177J_5 + 14653J_7 - 299J_9 = (0^{\circ}.89 \pm 2) \times 10^{-2}$$

and for the secular motions:

$$\dot{\omega} = 4499.379J_2 - 1267.4J_4 - 2537J_6 + 2117J_8 + 0^{\circ}.001901$$

$$+ 0^{\circ}.000568 = 4^{\circ}.87433 \pm 4$$

$$\dot{\Omega} = -3020.437J_2 + 2268.5J_4 - 95J_6 - 1036J_8 + 0^{\circ}.000221$$

$$- 0^{\circ}.000396 = -3.273988 \pm 7$$

The last two terms on the left hand side of the secular terms are the J_2^2 terms and lunisolar secular perturbations respectively.

In a 1962 report Kozai[14] used 40 equations of the amplitudes and 40 equations of the secular motions of 13 satellites to evaluate J_n up to the ninth order.

The two principal investigators in this field has been Kozai and D. C. King-Hele. Their most recent values shown on the next page represent an improvement over earlier ones because of the utilization of accurate Baker-Nunn camera tracking data.

Kaula believes the discrepancies between the two sets can be blamed either on the use of King-Hele of some lower inclination orbits of weak determination, Explorer 11 (i = 28.8°) and Explorer 9 (i = 38.8°), or on the use by Kozai of perigee motion as well as node motion for all satellites.

The chart of values is shown below. King-Hele[20] notes that if the potential is to be represented by fewer even harmonics, the best number is three and the recommended values are $J_2 = 1082.76 \times 10^{-6}$, $J_4 = -1.56 \times 10^{-6}$ and $J_6 = +0.39 \times 10^{-6}$.

All Values Times 10^{-6}

	Y. Kozai[21]	King-Hele[20]	NASA Recommended (1964)
J_2	1082.65 ± .12	1082.70 ± .10	1082.76
J_3	- 2.53 ± .02		- 2.55
J_4	- 1.62 ± .04	- 1.40 ± .20	- 1.56
J_5	- 0.21 ± .03		- 0.15
J_6	0.61 ± .08	0.37 ± 0.2	+ 0.39
J_7	- 0.32 ± .03		- .35
J_8	- 0.24 ± .11	0.07	0
J_9	- 0.10 ± .04		0
J_{10}	- 0.10 ± .12	- .50	0
J_{11}	+ 0.28 ± 0.04		0
J_{12}	- 0.28 ± .11	0.31	0
J_{13}	- 0.18 ± .04		0
J_{14}	+ 0.19 ± .13		0

What about higher order J_n terms; how much error do we have in the potential function by neglecting them? No one seems to have answered this satisfactorily. They must decrease faster than $\frac{1}{n}$ to insure convergence but there is always the possibility that a larger group, say

J_{50}-J_{75}, might be of the same sign and together contribute a significant error to the geoid. King-Hele[13] estimates they vary like $\frac{1}{n^2}$ and that the error in ignoring terms above J_8 is about 0.4×10^{-6}. Thus since neglected values readjust our determination of all values we have the problem that as each new J_n value is determined all the others have to be revised. Perhaps some better expansion than Legendre polynomials or expansion about a triaxial ellipsoid would help us. From a practical standpoint the uncertainties due to drag (because of density uncertainties) and other perturbations will limit the necessity for being too accurate in computing the potential function. In addition we shall soon reach some J_n term such that after each major earthquake all higher J_n terms will have to be revised. In deed, higher order J_n terms may well be time dependent.

Now that we have these J_2 and J_3 values it might be appropriate to make some remarks about what shape the symmetrical portion of the earth is in.

The J_2 term is by far the largest and most important term. It represents the flattening of the earth and corresponds to a flattening ratio of $\frac{1}{298.24}$. With a_e = 6378165 meters, this flattening ratio gives a polar axis of 6356779 meters giving use to a difference of 21.386 km. The presatellite difference was thought to be 21.467 km. This difference, 81 meters or 266 feet, seems negligible but it is important to geodesists who like measurements accurate to 10 meters or better and more important it proves the earth is not in hydrostatic equilibrium. This means that

the earth's interior has great strength of the order of 2×10^7 dynes/cm^3 which is required to support stresses at the base of the mantle, and the assumption that it can be treated as if it were a fluid, an assumption widely made in presatellite days, is simply not true.

The J_3 term shows the earth has a slight tendency, amounting to about 30 meters or 100 feet, towards a pear shape with the "stem" at the North Pole. Again 100 feet is pretty small but important to geophysicists because the earth could not assume this rather odd shape unless it had considerable internal strength. The pear problem also poses another question of why is the stem at the north rather than the south - a question yet to be answered.

The discovery of the pear-shaped earth came as a great surprise to geodesists of the 20th century. But it would have been no surprise to an eminent navigator of the 15th century - Christopher Columbus - who in the last years of his life stated that the earth "was not round in the way that is usually written, but that it has the shape of a pear that is very round, except in the place where the stem is, which is higher - ...[7]". How about that.

The fourth and sixth harmonics contribute only a little less than half the third and so they too make an important contribution to the earth's shape. All of these are small bumps indeed. On an 18 inch globe of the world the maximum difference would only be 1/16" and the highest mountains about 1/75". The world is much smoother and rounder than most bowling balls in the Fairborn bowling alley.

The figure on the next page shows the geoid relative to a spheroid of flattening of $\frac{1}{298.24}$ using earlier determinations of King-Hele of $J_2 = 1082.86 \times 10^{-6}$, $J_3 = -2.45 \times 10^{-6}$, $J_4 = -1.03 \times 10^{-6}$ and $J_6 = 0.72 \times 10^{-6}$, $J_5 = -0.05 \times 10^{-6}$.

Now in addition to these gravity perturbations there are also longitudinal dependent terms in the earth's potential field. In particular the equatorial plane of the earth is not a circle but more elliptical. This fact was suspected even before satellite data and some mention of the earth as a triaxial ellipsoid was heard. The table listed below gives a listing by Jung[1] of some gravity measurements with longitude terms. According to all these presatellite data the long axis of the equator was close to 0 degrees and 180 degrees longitude while the short axis is close to ± 90 degrees. The difference between the two equatorial semi-axes was between 130 and 350 meters.

LONGITUDE TERMS FOR TRIAXIAL ELLIPSOID[1]

Author	Year	$a_1 - a_2$ (Meters)	Longitude
Helment	1915	230	17°W
Berroth	1916	200	10°W
Heiskanen	1924	340	18°E
Heiskanen	1928	240	0°
Heiskanen	1938	350	25°W
Jeffreys	1942	160	0°
Jung	1940	370	25°W
Jung	1941	200	10°W
Niskanen	1945	290	4°W
Jung	1948	130	10°W

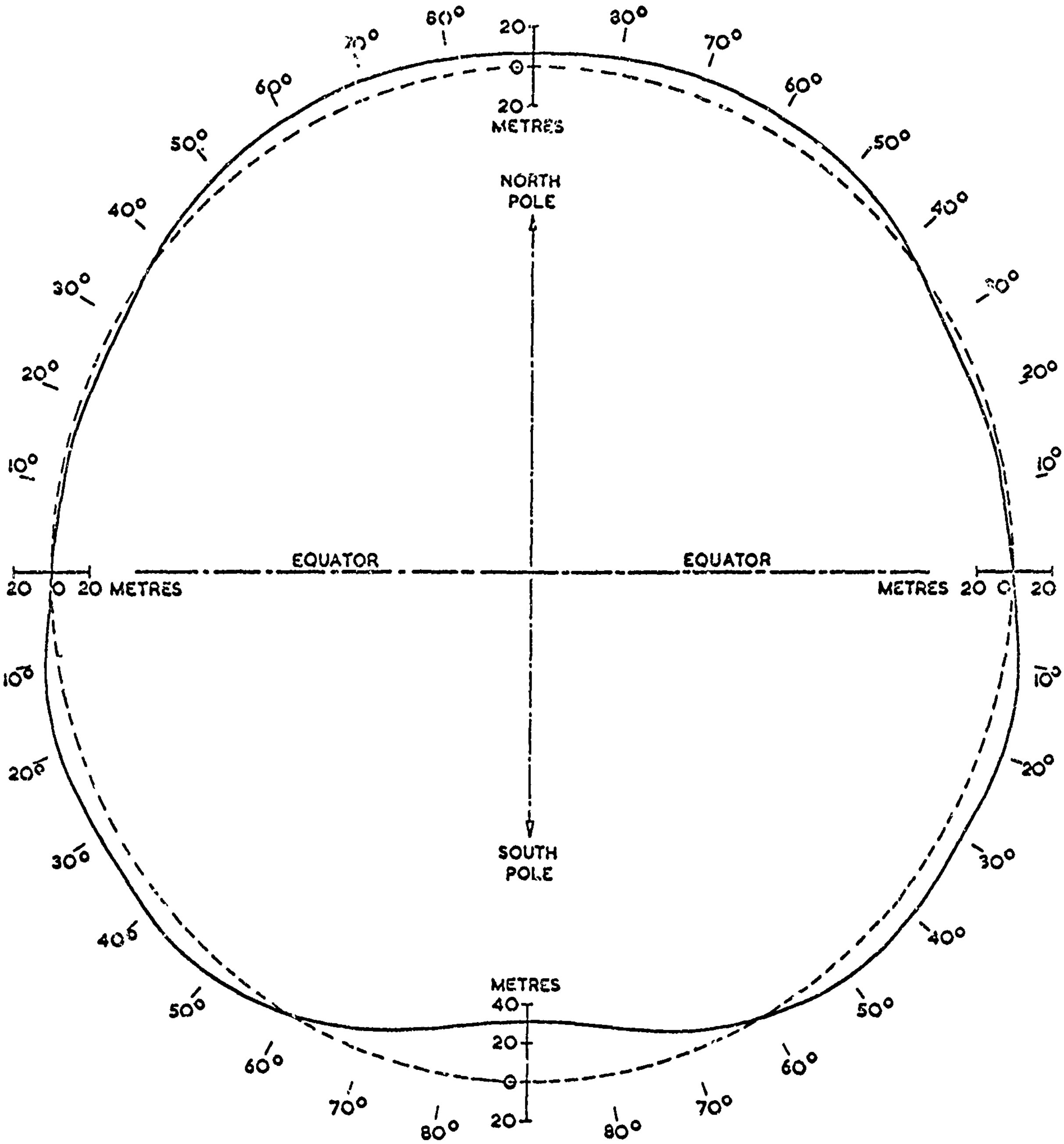


FIGURE 6
THE HEIGHT IN METRES OF THE GEOID (SOLID LINE) RELATIVE TO A SPHEROID OF FLATTENING 1/298·24 (BROKEN LINE), AS GIVEN BY THE ZONAL HARMONICS OF KING-HELE.

Satellites tend to sample all longitudes impartially as the earth spins under them, and are therefore not much affected by slight variations in longitudinal dependent gravity forces. However, these variations do cause small but characteristic perturbations from which the longitude dependent terms can be determined. Recall J_{22} affects the 24 hour satellite and thus offers a technique for its measurement.

The shape of the earth's equator is not exactly an ellipse but a more complex figure built of many individual harmonics in precisely the same manner as we developed the symmetrical J_n terms. The complete geopotential function can be represented in the form

$$U = \left[\frac{\mu}{r}\, 1 - \sum_{n=2}^{n_1} J_n \left(\frac{a_e}{r}\right)^n P_n(\sin L) + \sum_{n=2}^{n_2} \sum_{m=1}^{n} J_{nm} \left(\frac{a_e}{r}\right)^n P_n^m (\sin L) \cos m(\lambda - \lambda_{nm})\right]$$

$\mu = k^2 M_e$, r, L, λ are spherical coordinates of radius, latitude and east longitude of the field point, a_e is equatorial mean radius, P_n are Legendre polynomials of degree n, P_n^m or as it is sometimes written $P_{n,m}$ are associated Legendre functions of the first kind of degree n and order m.; J_n and J_{nm} are numerical constants, λ_{nm} is the longitude associated with J_{nm} and n_1, n_2 are maximum degrees of the zonal and tesseral harmonics included in the computation of U. The associated Legendre functions can be found from -

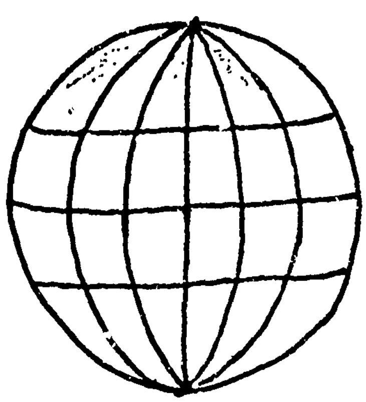
$$P_n^m (\sin L) = (1 - \sin^2 L)^{m/2} \frac{d^m}{d(\sin L)^m} [P_n (\sin L)]$$

$$= (1 - \sin^2 L)^{m/2} \frac{1}{2^n m!} \frac{d^{n+m}}{d(\sin L)^{n+m}} (\sin^2 L - 1)^n$$

$$P_n^m (\sin L) = (1 - \sin^2 L)^{m/2} \, 2^m m! \, P_{n+m} (\sin L)$$

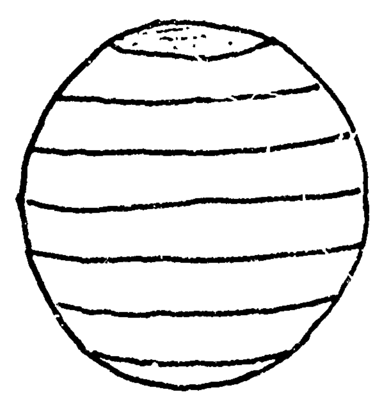
When these associated Legendre functions are multiplied by $\sin m\lambda$ or $\cos m\lambda$ they become tesseral harmonics. The figure below indicates a tesseral harmonic where the shaded area represents tesserae where the function is positive and the unshaded area represents negative tesserae. When $m = 0$, tesseral harmonics divide the sphere into zones of alternate positive and negative values - hence are called zonal harmonics. These arise with the J_n symmetrical terms.

TESSERAL



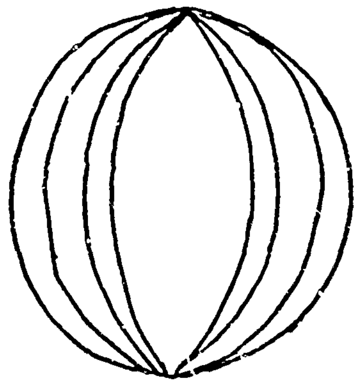
$P_{9,6}$ (sin L) sin 6λ

ZONAL



$P_{7,0}$ (sin L)

SECTORIAL



$P_{7,7}$ (sin L) $\begin{Bmatrix} \cos 7\lambda \\ \sin 7\lambda \end{Bmatrix}$

When m = n the tesseral harmonics divide the sphere into sectors of alternate positive and negative values - hence are called sectorial harmonics.

Another form of the geopotential has been adopted by the International Astronomical Union[2], unfortunately, and both forms abound in the literature.

$$U = \frac{\mu}{r}\left[1 + \sum_{n=2}^{\infty} \sum_{m=0}^{n} \left(\frac{a_e}{r}\right)^n P_n^m (\sin L) \; (C_{nm} \cos m\lambda + S_{nm} \sin m\lambda)\right]$$

And variations occur with this formula! Some people use a different normalization for the associated Legendre polynomials or spherical harmonics as they are called.

$$P_n^m (\sin L) = P_{nm} (\sin L) = (1 - \sin^2 L)^{m/2} \frac{1}{2^n m!} \frac{d^{n+m}}{d(\sin L)^{n+m}} (\sin^2 L - 1)^n$$

We can expand the last term in a series to write

$$P_{nm}(\sin L) = \frac{(1 - \sin^2 L)^{m/2}}{2^n m!} \frac{d^{n+m}}{d(\sin L)^{n+m}} \left[\sum_{t=0}^{n} \frac{n!(-1)^t}{(n - t)!t!} (\sin L)^{2n-2t} \right]$$

This may be differentiated to give

$$P_{nm}(\sin L) = \frac{(1 - \sin^2 L)^{m/2}}{2^n} \sum_{t=0}^{n} \frac{(2n - 2t)!\,(-1)^t (\sin L)^{n-m-2t}}{(n - m - 2t)!\,(n - t)!\,t!} \quad .$$

This form of the spherical harmonic is used in the term

$$\sum_{n=2}^{\infty} \sum_{m=1}^{n} J_{mn} \left(\frac{a_e}{r}\right)^n P_{nm}(\sin L) \cos m(\lambda - \lambda_{nm})$$

and in the term

$$\sum_{n=2}^{\infty} \sum_{m=0}^{n} \left(\frac{a_e}{r}\right)^n P_{nm}(\sin L) \left[C_{nm} \cos m\lambda + S_{nm} \sin m\lambda\right] .$$

However, Kaula and others use the "fully-normalized" spherical harmonics which are given by $\left[\begin{array}{ll} \delta_{m0} = 1 & \text{for } m = 0 \\ \delta_{m0} = 0 & \text{for } m \neq 0 \end{array}\right]$

$$\overline{P}_{nm} (\sin L) = \sqrt{\frac{(n - m)!(2n + 1)(2 - \delta_{om})}{(n + m)!}} (1 - \sin^2 L)^{m/2} \sum_{t=0}^{k}$$

$$\frac{(-1)^t (2n - 2t)!(\sin L)^{n - m - 2t}}{2^n \; t! \; (n - t)!(n - m - 2t)!}$$

where k is the integer part of $\frac{n - m}{2}$ and δ_{om} is the Kronecker delta functions. The corresponding potential term in this case is usually written

$$U = \frac{\mu}{r}\left[1 + \sum_{n=2}^{\infty} \sum_{m=0}^{n} \left(\frac{a_e}{r}\right)^n \overline{P}_{nm} (\sin L) \left[\overline{C}_{nm} \cos m \lambda + \overline{S}_{nm} \sin m \lambda\right]\right] .$$

For this case $\overline{C}_{no} = - \frac{J_n}{\sqrt{2n + 1}}$; for the other potential function $C_{no} = J_n$ and in fact

$$C_{no} = - J_n \qquad\qquad S_{no} = 0$$

$$C_{nm} = J_{nm} \cos m \lambda_{nm}$$

$$S_{nm} = J_{nm} \sin m \lambda_{nm}$$

$$J_{nm} = \sqrt{C_{nm}^2 + S_{nm}^2}$$

$$\lambda_{nm} = \frac{1}{m} \arctan \frac{S_{nm}}{C_{nm}}$$

(take smallest value)

and

$$\left\{ \bar{C}_{nm}, \bar{S}_{nm} \right\} = \sqrt{\frac{(n-m)!}{(n+m)!}} \left\{ C_{nm}, S_{nm} \right\} .$$

So one must beware of geophysicists bearing C and **S**, which is probably why one often hears the expression, "those **CS** geophysicists".

Again we can expand the potential function and integrate term by term and identify these new J_{nm} coefficients as we did with the J_n coefficients. The values for the various C_{nm} or S_{nm} or J_{nm} tesseral harmonic coefficients are not very accurately known.

Mercer[22] gives the following list of the best results from the TRANSIT Program.[23] In making these determinations, APL uses King-Hele's[20] values for the even zonal harmonics and Newton's[15] (no relation) value for the odd degree zonals. They use $a_e = 6378166$ m and $\mu = k^2 M_e = 3.986075 \times 10^{14}$ m^3/sec^2. This list appears on the next page.

The various determinations of these constants by different investigators seem to give rise to a geoid which has basically the same general

phase, i.e., the relative maximum and minimum occur in the same general location, but the magnitudes differ by as much as 30 meters. The APL values listed in the table have fewer possible sources of error but will most certainly be improved with age as longer arc lengths are tracked and camera data is integrated with the doppler information.

TESSERAL COEFFICIENTS

In units of 10^{-6} and degrees east longitude.

n	m	J_{nm}	λ_{nm}	C_{nm}	S_{nm}
2	1	0	0	0	0
2	2	1.723	- 13.18	1.544	-0.765
3	1	1.856	5.31	1.848	0.172
3	2	0.4482	- 13.33	0.4006	-0.2011
3	3	0.1643	16.00	0.1100	0.1221
4	1	0.7278	-136.68	- .5295	- .4993
4	2	0.1536	27.47	0.0882	0.1257
4	3	.05719	- 1.10	.05709	- .0033
4	4	.004781	36.01	- .00387	.00281
5	1	0.1509	- 81.86	.0214	- .1493
5	2	0.05329	- 2.60	.05307	- .00483
5	3	0.005843	- 3.45	.005747	- .001049
5	4	0.004093	58.25	- .002463	- .003269
5	5	0.0008462	- 15.42	0.0001963	- .0008231
6	1	0.06611	157.21	- .06095	.02561
6	2	0.03746	113.15	- .02588	- .02708

n	m	J_{nm}	λ_{nm}	C_{nm}	S_{nm}
6	3	0.01166	- 1.75	.01161	- .00107
6	4	0.002256	60.74	- .001026	- .002010
6	5	0.0003328	- 17.75	.0000071	- .0003327
6	6	0.00005153	- 14.34	.00000356	- .00005140

From TRANSIT Program as reported by Mercer.[22]

Which model to use for U and how many coefficients to use depends on the inclination and period of the satellite. As was mentioned earlier in the notes, resonance can make certain of the J_{nm} terms become quite important in predicting that particular satellite's position. By this means we have the following:

$J_{13,13} = 0.52 \times 10^{-6}$ $\qquad \lambda_{13,13} = 10.4°$

$J_{14,14} = 0.56 \times 10^{-6}$ $\qquad \lambda_{14,14} = 15°$

$J_{15,14} = 0.08 \times 10^{-6}$ $\qquad \lambda_{15,14} = 19.6°$

[Yionoulis, Dec 1965][39]

Wagner in 1964[25] said the Syncom satellite showed $J_{22} = -(1.70 \pm .05) \times 10^{-6}$ $\quad \lambda_{22} = -(19 \pm 6)°$ (West Longitude). In 1965 he corrected for higher order J_{nm} terms and obtained a value of[26]

$$J_{22} = -(1.9 \pm 0.2) \times 10^{-6} \qquad \lambda_{22} = -(21 \pm 7)^{\circ}$$

This latter determination corresponds to a difference between major and minor semi-axis of 73 $\pm$ 8 meters in the equatorial plane. Previous (1964) determination led to 65 $\pm$ 2 meters difference. To obtain an idea of the accuracy available, several of the more recent determinations of tesseral coefficients are listed on the next page. The data are from Wagner's article.[26]

A better understanding of the effect of each of the potential terms can be had by representing the tesseral harmonics on a sphere. Wagner's pictures are shown on the next two pages. The shaded areas are positive terms but with continuously varying values, the numbers listed on the figures are the maximum deviations from the sphere.

To be more realistic we need to take all of these tesseral harmonics and form a geoid map. Which geoid picture we obtain depends on which oblate spheroid we select as the reference ellipsoid and, of course, whose $J_{n,m}$ coefficients we use. The resultant picture will show deviations of the equipotential surface, the geoid, from the selected reference ellipsoid. Several such geoid maps were developed by Kaula[29] using an ellipsoid with $1/f = 298.24$ and $a_e = 6378165$ meters. These geoid maps are shown on the following pages.

The important features are the depression, some 40-70 meters deep, south of India and the elevation, some 20 to 60 meters high, near New Guinea; these are similar to the main features of the simple picture of an elliptic equator; however other features of this simple

Table A1

TESSERAL COEFFICIENTS IN THE EARTH'S GRAVITY POTENTIAL $\left\{ V_E = \frac{\mu}{r} \sum_{n=2}^{\infty} \sum_{m=0}^{n} \left[1 - (R_0/r)^n P_n^m (\sin \phi) J_{nm} \cos m (\lambda - \lambda_{nm}) \right] \right\}$ AS REPORTED 1859-1964*

TESSERAL GEOID†	J_{22}	λ_{22}	J_{31}	λ_{31}	J_{32}	λ_{32}	J_{33}	λ_{33}	J_{41}	λ_{41}	J_{42}	λ_{42}	J_{43}	λ_{43}	J_{44}	λ_{44}
(1) Wagner (1964)[A]	-1.7×10^{-6}	-19.0°														
(2) Kaula-Combined (1964)[E]	-1.51	-15.5	-1.51×10^{-6}	0.0°	-0.102×10^{-6}	0.0°	-0.149×10^{-6}	22.8°	-0.465×10^{-6}	-136.0°	-0.163×10^{-6}	37.0°	-0.061×10^{-6}	- 1.9°	-0.0053×10^{-6}	35.8°
(3) Izsak (1964)[B]	-1.00	-17.0	-0.934	-15.5	-0.116	19.0	-0.173	38.0	-0.449	-146.0	-0.074	47.5	-0.024	- 3.9	-0.0206	23.3
(4) Kaula (1964)[B]	-1.77	-18.2	-2.12	- 5.4	-0.379	10.5	-0.105	23.1	-0.263	-239.0	-0.117	42.3	-0.0473	15.0	-0.0104	14.5
(5) Anderle and Oesterwinter (1963)[A]	-2.09	-14.1							-0.773	-140.6	-0.287	34.7	-0.163	- 4.3		
(6) Guier (1963)[A]	-1.80	-10.4	-1.77	6.3	-0.286	- 2.6	-0.204	24.1	-0.73	-141.0	-0.273	38.6	-0.0791	- 0.7	-0.0102	35.0
(7) Kaula (Sept. 1963)[B]	-1.51	-18.1	-1.65	5.3	-0.141	46.4	-0.145	15.8	-0.471	-228.0	-0.078	44.2	-0.0265	22.6	-0.0035	23.3
(8) Izsak (July 1963)[B]	-1.05	-11.2	-1.1	3.2	-0.20	-21.8	-0.14	20.0	-0.43	-132.1	-0.13	37.0	-0.026	13.5	-0.019	14.8
(9) Kaula (May 1963)[B]	-1.4	-21.5	-1.6	- 1.9	-0.15	35.8	-0.156	18.5	-0.53	-233.7	-0.12	44.5	-0.019	10.7	-0.0039	23.3
(10) Cohen (May 1963)[A]	-2.08	-14.1							-0.775	-159.0	-0.258	34.6	-0.162	- 4.3		
(11) Kaula (Jan. 1963)[B]	-1.62	-21.4	-1.81	- 3.57	-0.145	6.6	-0.112	37.6	-0.179	-245.5	-0.072	47.7	-0.0935	5.9	-0.0132	28.4
(12) Uotila (1962)[C]	-1.52	-36.5	-0.685	-81.0	-0.409	- 5.2	-0.398	19.5	-0.238	-127.0	-0.211	14.6	-0.032	- 9.3	-0.0142	- 2.6
(13) Kozai (Oct. 1962)[B]	-1.2	-21.[illegible]	-1.2	4.6	-0.14	-16.8	-0.10	42.6	-0.52	-122.5	-0.062	65.2	-0.035	0.5	-0.031	14.9
(14) Newton (April 1962)[A]	-2.15	-10.9							-2.53	-139.1						
(15) Newton (Jan. 1962)[A]	-4.16	-11.0														
(16) Kozai (June 1961)[B]	-2.32	-37.5	-3.21	22.0	-0.41	31.0	-1.91	51.3	-0.262	-196.5	-0.168	54.0	-0.014	-13.0	-0.054	50.3
(17) Kaula (June 1961)[E]	-0.55	-13.3	-1.19	20.6	-0.33	- 0.9	-0.21	22.6	-0.617	-166.0	-0.14	21.1	-0.031	- 0.5	-0.003	26.4
(18) Izsak (Jan. 1961)[B]	-5.35	-33.2														
(19) Kaula (1961)[C]	-1.68	-38.5							-1.15	- 13.0						
(20) Kaula (1959)[C]	-0.62	-20.9	-0.98	55.4	-0.11	13.3	-0.19	14.3	-0.46	-132.3	-0.081	48.6	[illegible]	-30.0	-0.02	22.5
(21) Jeffreys (1959)[C]	-4.17	0.0														
(22) Uotila (1957)[C]	-3.5	- 6.0														
(23) Zhongolovitch (1957)[C]	-5.95	- 7.7	-2.21	-25.7	-0.628	-26.4	-0.54	13.0	-0.78	-149.1	-0.030	45.0	-0.051	- 3.8	-0.0224	15.9
(24) Subbotin (1949)[C]	-5.5															
(25) Lambert (1915)[C]	-9.2	-25.0														
(26) Niskanen (1915)[C]	-7.67	- 4.0														
(27) Krassowski (1912)[D]	-5.6	15.0														
(28) Jeffreys (1912)[C]	-4.1	0.0	-2.1	0.0	-0.66	5.0	-0.24	33.3								
(29) Heiskanen (1938)[C]	-9.2	-25.0														
(30) Hirvonen (1934)[C]	-3.6	-19.0														
(31) Heiskanen (1929)[C]	-4.3	38.0														
(32) Heiskanen (1923)[C]	-6.34	0.0														
(33) Heiskanen (1924)[C]	-9.0	18.0														
(34) Berroth (1916)[C]	-3.9	-10.0														
(35) Helmert (1915)[C]	-6.0	-17.0														
(36) Hill (1881)[D]	-32.9	17.0														
(37) Clarke (1878)[D]	-12.1	- 8.0														
(38) Clarke (1866)[D]	-50.9	15.0														
(39) Clarke (1860)[D]	-42.3	14.0														
(40) Schubert (1859)[E]	-18.8	41.0														

*(r is the radial distance of the field point to the center of mass of the earth, μ the earth's Gaussian gravity constant $\doteq 3.9860 \times 10^{20} cm^3/sec^2$, R_0 the mean equatorial radius of the earth $\doteq$ 6378.2 km; ϕ is the geocentric latitude of the field point, λ the geographic longitude of the field point; $J_{21} = 0$, since the polar axis is very nearly a principal axis of inertia for the earth. $P_n^m(\sin \phi) = \cos^m \phi \sum_{t=0}^{K} T_{nmt} \sin^{n-m-2t} \phi$, where K is the integer part of $(n - m)/2$ and $T_{nmt} = (-1)^t (2n - 2t)!/2^n t! (n - t)! (n - m - 2t)!$ (see footnote, p. 32); the tesseral coefficients are those for which $(m \neq 0)$.

The J_{nm}'s and λ_{nm}'s in this table, except in one or two instances, have been converted from the original authors' set of gravity coefficients. The blanks indicate the author did not consider that particular harmonic in fitting an earth potential to the observed data. In one or two instances (noted below) the author reported tesseral coefficients to higher order than the fourth.

†The second page of Table A1 consists of a list of references and comments for geoids (1)-(40). The superscript letters A-E in this column indicate the type of geoid as given below:

[A]Satellite-Doppler geoid

[B]Satellite-camera geoid

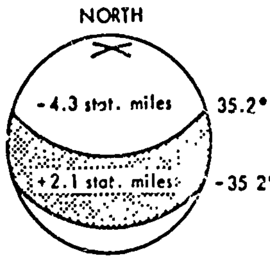
[C]Surface-gravimetric geoid

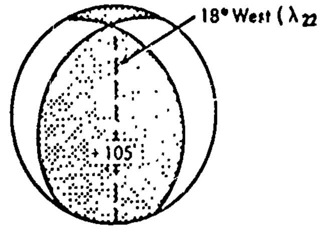
[D]Astrogeodetic geoid

[E]Combined astrogeodetic, gravimetric, and satellite geoid

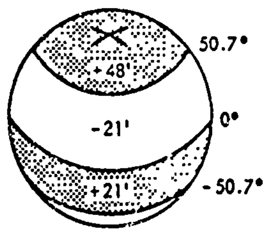
Table A1 (Continued)

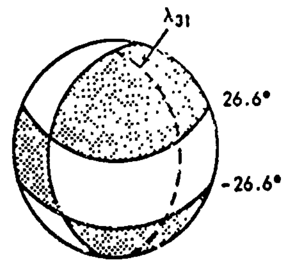
GEOID	REFERENCES	COMMENTS
(1)	"Determination of the Triaxiality of the Earth from Observations on the Drift of the Syncom II Satellite," NASA Goddard Space Flight Document X-621-64-90, April 1964.	Uses 7 months of Doppler data during longitude drift over Brazil
(2)	Private communication from W. M. Kaula, July 1964.	A weighted average of geoids (3),(4),(5),(6),(12),(13), and (17)
(3),(4)	Private communication from W. M. Kaula, July 1964.	Geoid (3) is a revision of (8) with new data, (4) is a revision of (7) with new data
(5)	Private communication from W. M. Kaula, July 1964.	Appears to be a revision of geoid (10)
(6),(10),(14)	In. "Non-Zonal Harmonic Coefficients of the Geopotential from Satellite Doppler Data," W. H. Guier and R. H. Newton, Johns Hopkins Univ. APL-TG-520, Nov. 1963	
(7)	"Improved Geodetic Results from Camera Observations of Satellites," J. Geophys. Res. 68(18), Sept. 1963.	Uses five satellites, of medium and high inclination
(8)	"Tesseral Harmonics in the Geopotential," Nature, pp. 137-139 July 13, 1963.	Uses five satellites, of medium and high inclination.
(9)	Private communication from W. M. Kaula, May 1963.	Uses three satellites of medium inclination
(11)	"Tesseral Harmonics of the Gravitational Field and Datum Shifts Derived from Camera Observations of Satellites," J. Geophys. Res. 68(2), Jan. 15, 1963.	Uses three satellites of medium inclination.
(12)	Private Communication from W. M. Kaula, July 1964.	
(13)	Private communication from Y. Kozai, Oct. 1962.	
(14)	Private communication from Newton, April 1962.	
(15)	"Ellipticity of the Earth's Equator Deduced from the Motion of Transit 4A," J. Geophys. Res. 67(1), Jan. 1962.	
(16)	"Tesseral Harmonics of the Gravitational Potential of the Earth as Derived from Satellite Motions," Astronom. J. 66(7), Sept. 1961.	Uses three satellites of medium inclination.
(17)	"A Geoid and World Geodetic System Based on a Combination of Gravimetric, Astrogeodetic, and Satellite Data," J. Geophys. Res. 66(6):1807, 1961.	Considers tesserals to eighth order.
(18)	From: Astronom. J. 66(5):226-229, June 1961.	Uses two satellites of medium inclination.
(19)	In: Space Research II, pp. 360-372, H. C. Van De Hulst, Ed: North Holland Publishing Co., Amsterdam.	
(20)	In: J. Geophys. Res. 64:2401, 1959.	
(21)	"The Earth," Cambridge Univ. Press: N.Y., 4th edition, chapt. IV, V.	
(22),(26),(32),(33),(35)	"The Earth and its Gravity Field," McGraw-Hill: N.Y., p. 79, 1958.	
(23)	In: Bull. Inst. Theoret. Astron. 6.505.	
(24)	"Course in Celestial Mechanics III," Gostehizdat: Moscow, p. 278, 1958.	
(27)	In: "Passive Dynamics in Space Flight," Bureau of Naval Weapons paper by J. D. Nicolaides and M. M. Macomber, 1962.	
(28),(27)	In: Monthly Notices Royal Astronom. Soc. Geophys. Suppl., 5(55), 1942.	
(27),(25),(29)-(31),(34),(36)-(40)	In: "Is the Earth a Triaxial Ellipsoid?" W. A. Heiskanen, J. Geophys. Res. 67(1), Jan. 1962.	



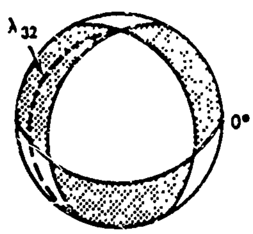


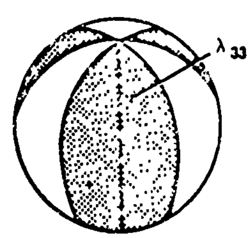
$$V_E = \frac{\mu_E}{r}\left[1 - \frac{J_{20}R_0^2}{2r^2}\left(3\sin^2\phi - 1\right) - 3J_{22}\frac{R_0^2}{r^2}\cos^2\phi\cos 2\left(\lambda - \lambda_{22}\right)\right.$$





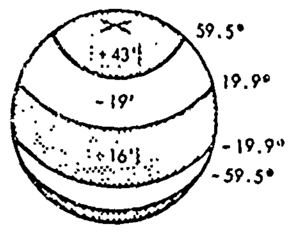
$$-\frac{J_{30}R_0^3}{2r^3}\left(5\sin^3\phi - 3\sin\phi\right) - \frac{J_{31}R_0^3}{2r^3}\cos\phi\left(15\sin^2\phi - 3\right)\cos\left(\lambda - \lambda_{31}\right)$$

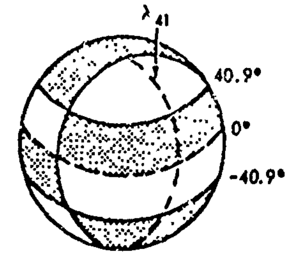




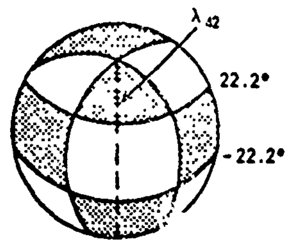
$$-15J_{32}\frac{R_0^3}{r^3}\cos^2\phi\sin\phi\cos 2\left(\lambda - \lambda_{32}\right) - 15J_{33}\frac{R_0^3}{r^3}\cos^3\phi\cos 3\left(\lambda - \lambda_{33}\right)$$

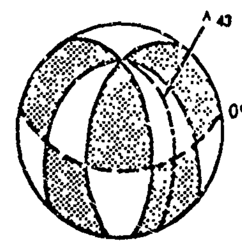
FIGURE 7



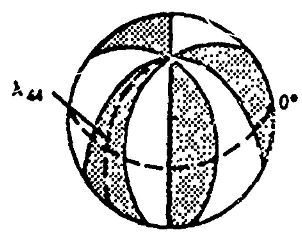


$$-\frac{J_{40} R_0^4}{8r^4}\left(35\sin^4\phi - 30\sin^2\phi + 3\right) - \frac{J_{41} R_0^4}{8r^4}\left(140\sin^3\phi - 60\sin\phi\right)\cos\phi\cos\left(\lambda - \lambda_{41}\right)$$





$$-\frac{J_{42} R_0^4}{8r^4}\left(420\sin^2\phi - 60\right)\cos^2\phi\cos 2\left(\lambda - \lambda_{42}\right) - \frac{J_{43} R_0^4}{8r^4}\, 840\sin\phi\cos^3\phi\cos 3\left(\lambda - \lambda_{43}\right)$$



$$-\frac{J_{44} R_0^4}{8r^4}\, 840\cos^4\phi\cos 4\left(\lambda - \lambda_{44}\right)\Big]$$

FIGURE 8

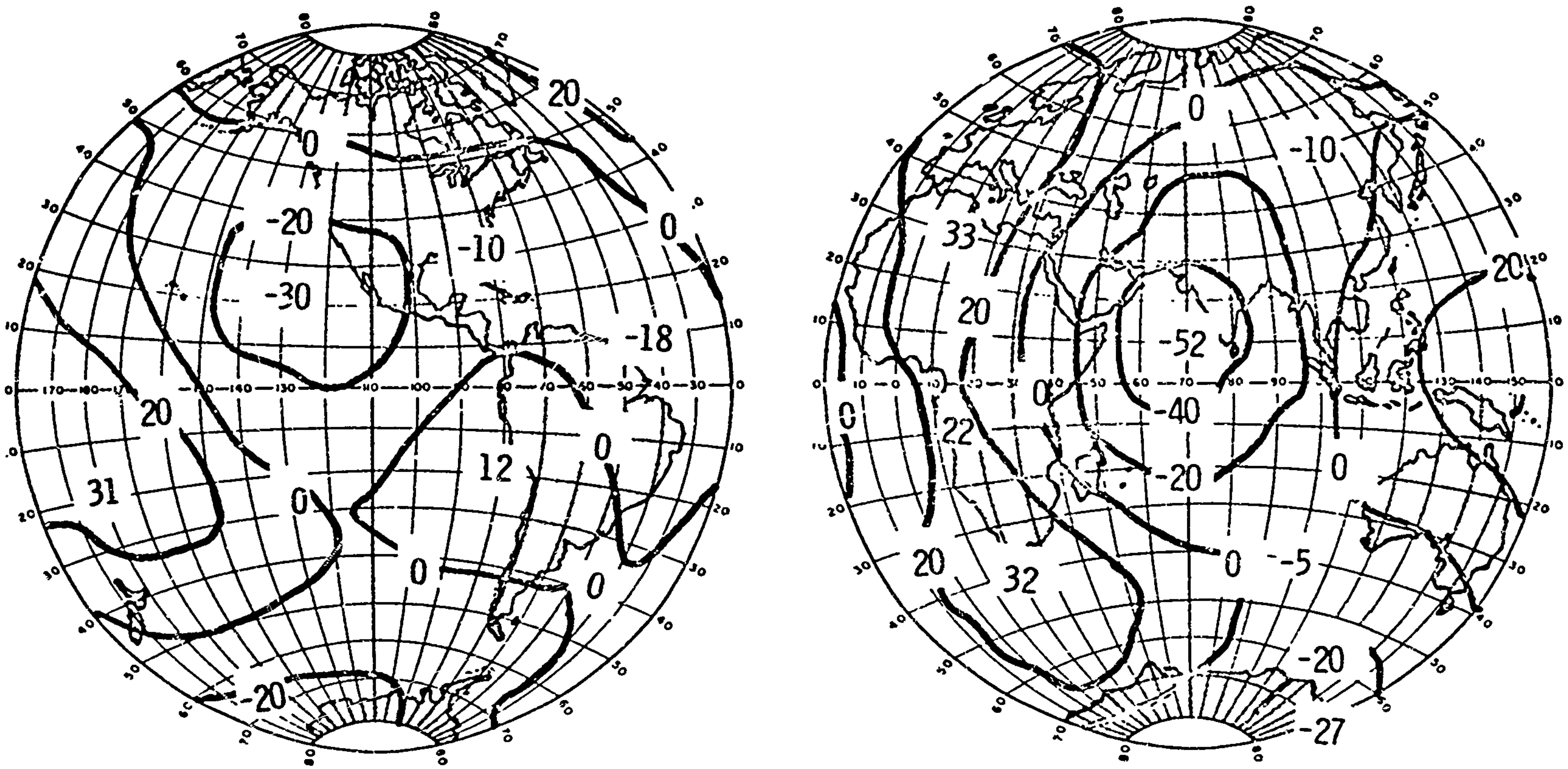


Figure 9. Geoid Heights in Meters Based on Harmonic Coefficients Determined from Camera Tracking of Satellites by Izsak (1963)

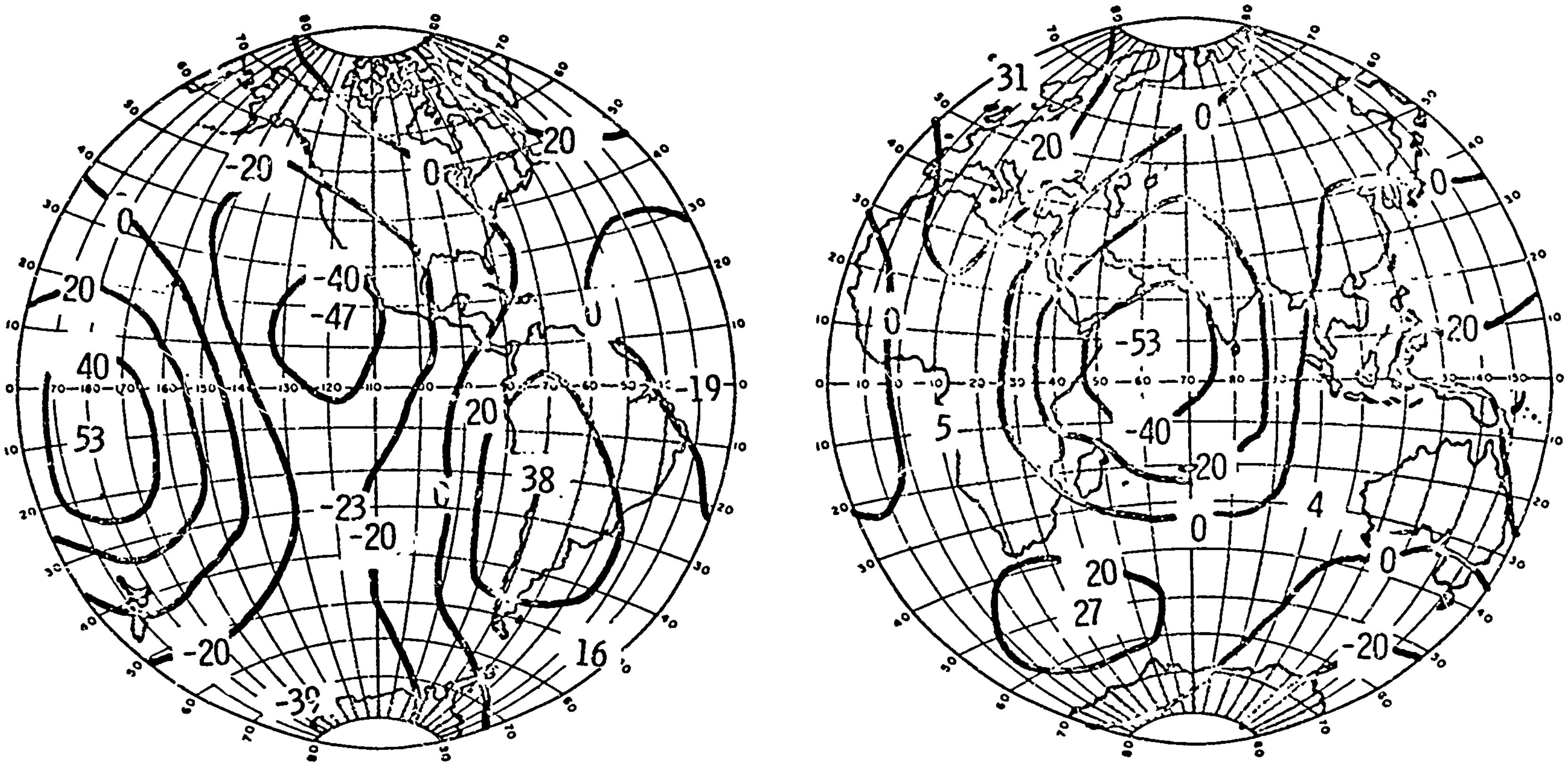


Figure 10. Geoid Heights in Meters Based on Harmonic Coefficients Determined from Camera Tracking of Satellites by Kozai (1963)

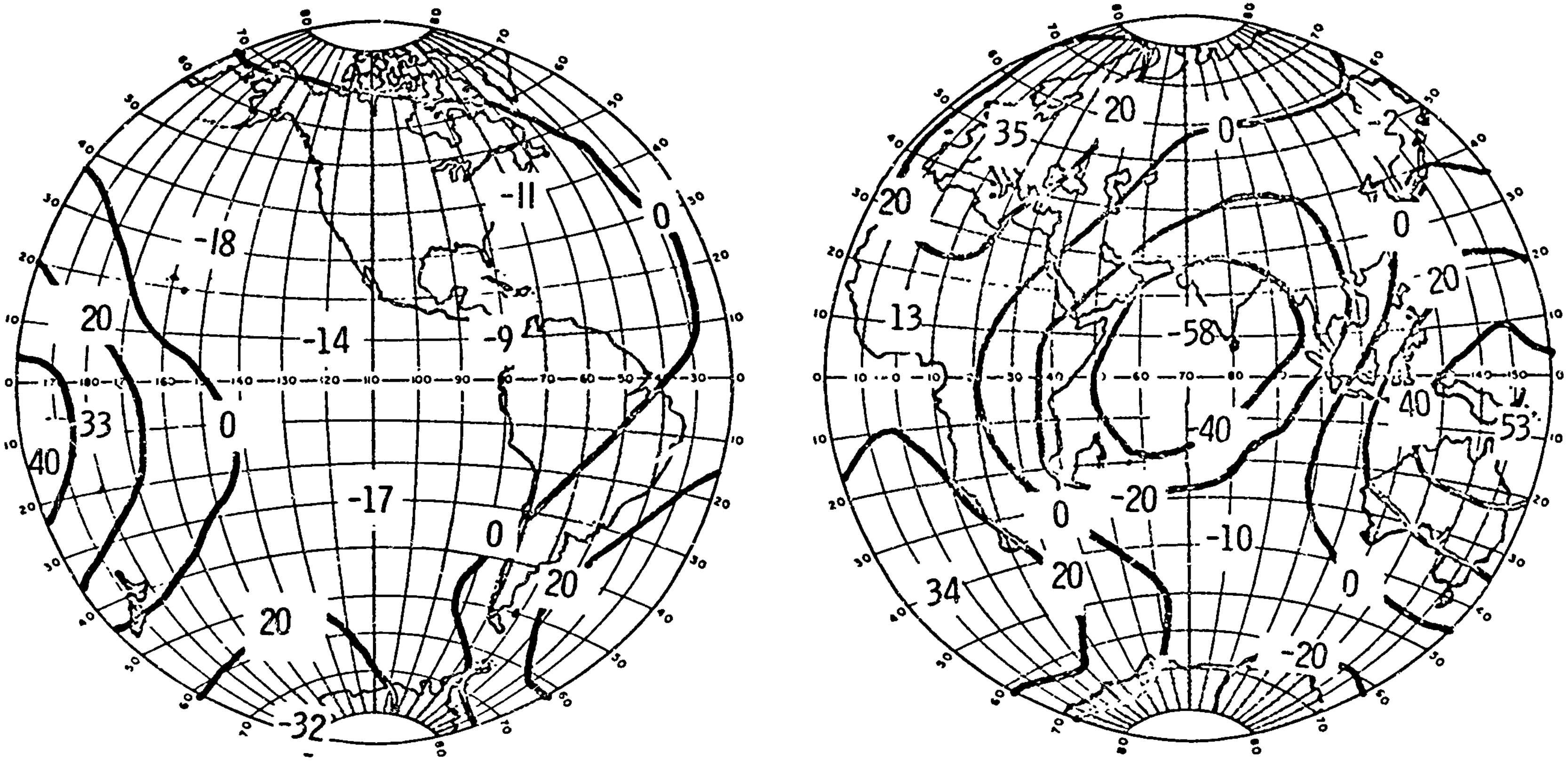


Figure 11. Geoid Heights in Meters Based on Harmonic Coefficients Determined from Camera Tracking of Satellites by Kaula (1963)

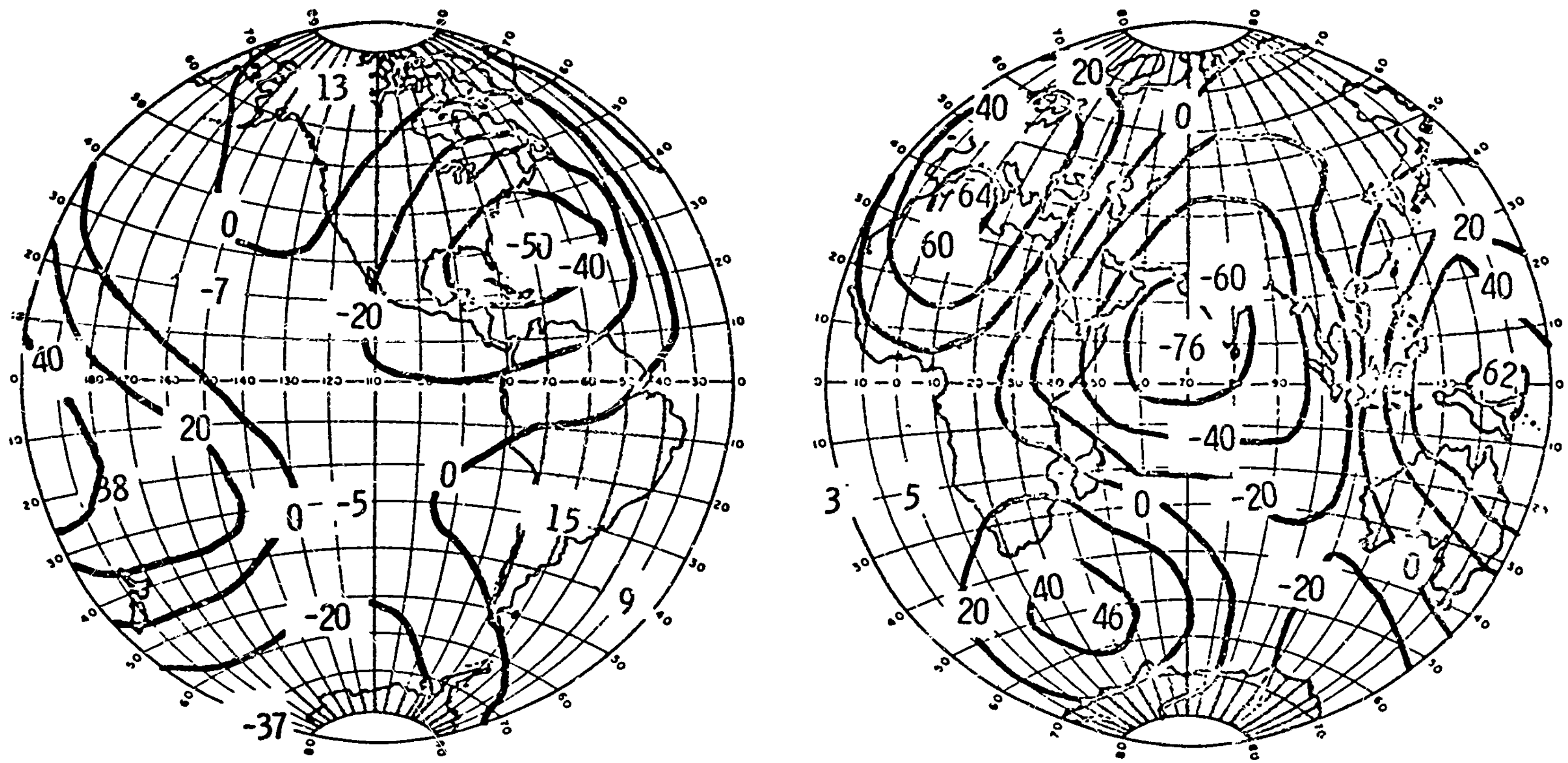


Figure 12. Geoid Heights in Meters Based on Harmonic Coefficients Determined from Doppler Tracking of Satellites by Guier (1963)

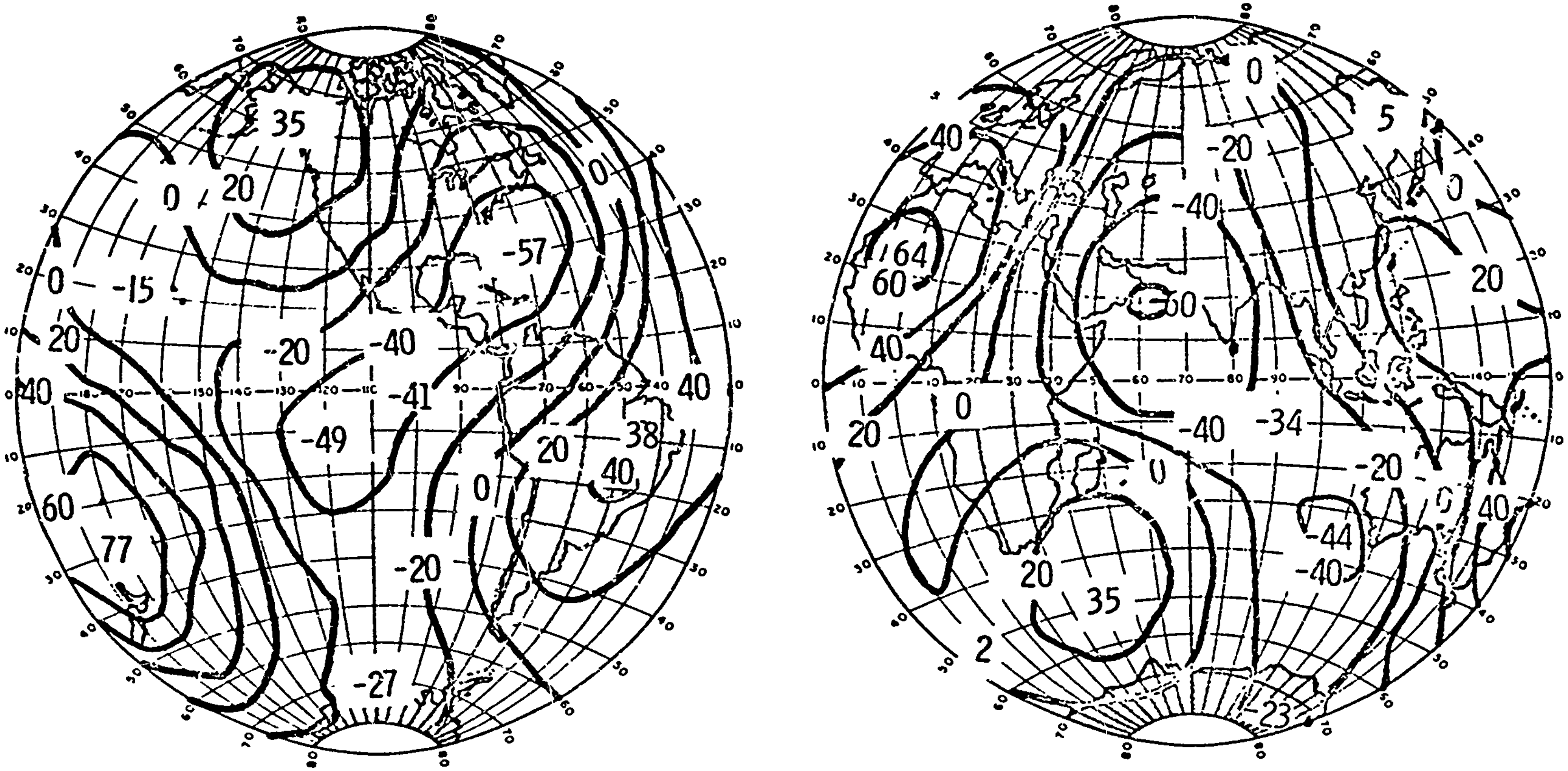


Figure 13. Geoid Heights in Meters Based on Harmonic Coefficients Determined from Doppler Tracking of Satellites by Anderle and Oesterwinter (1963)

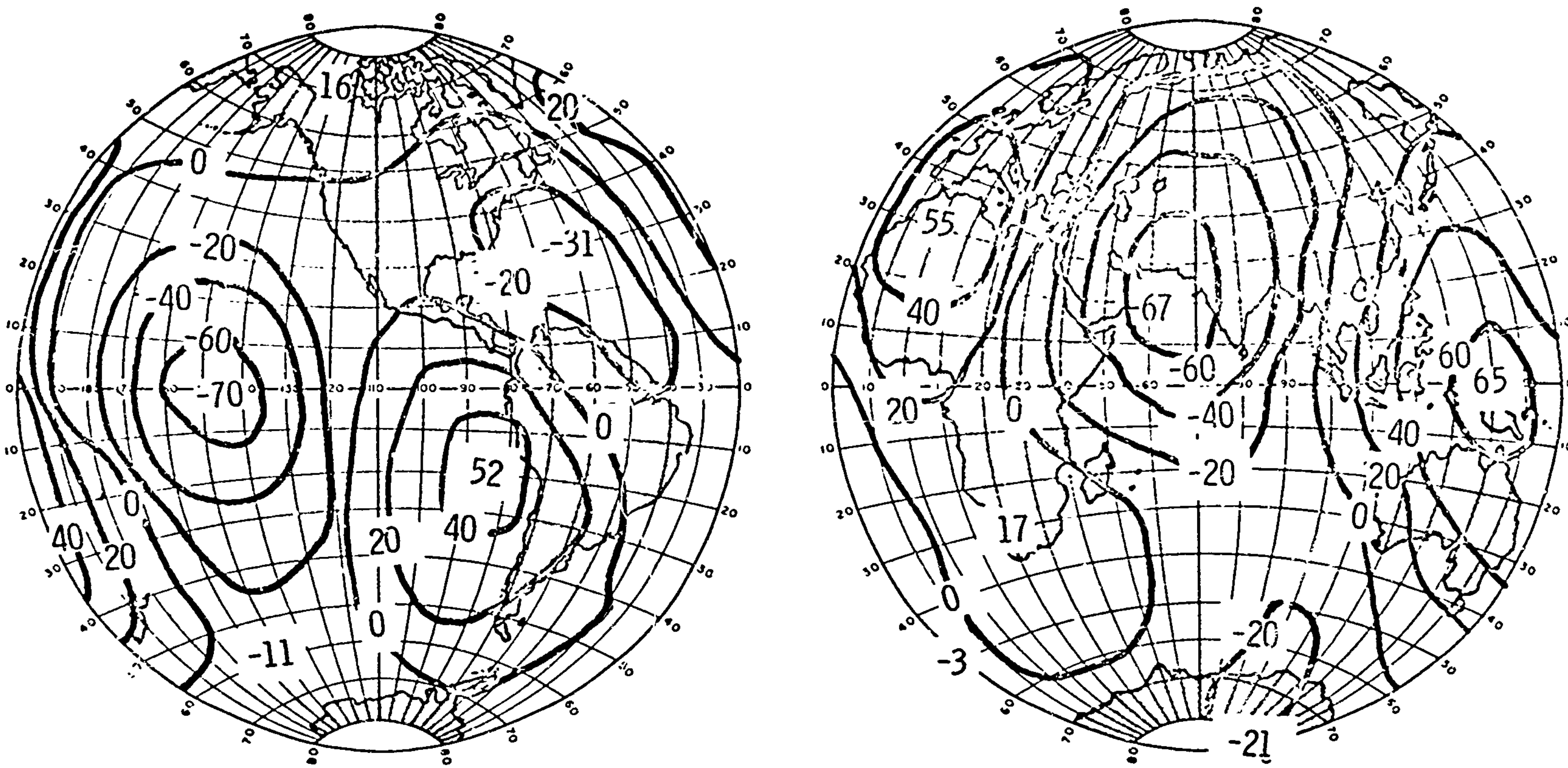


Figure 14. Geoid Heights in Meters Based on Harmonic Coefficients Determined from Terrestrial Gravimetry by Uotila (1962)

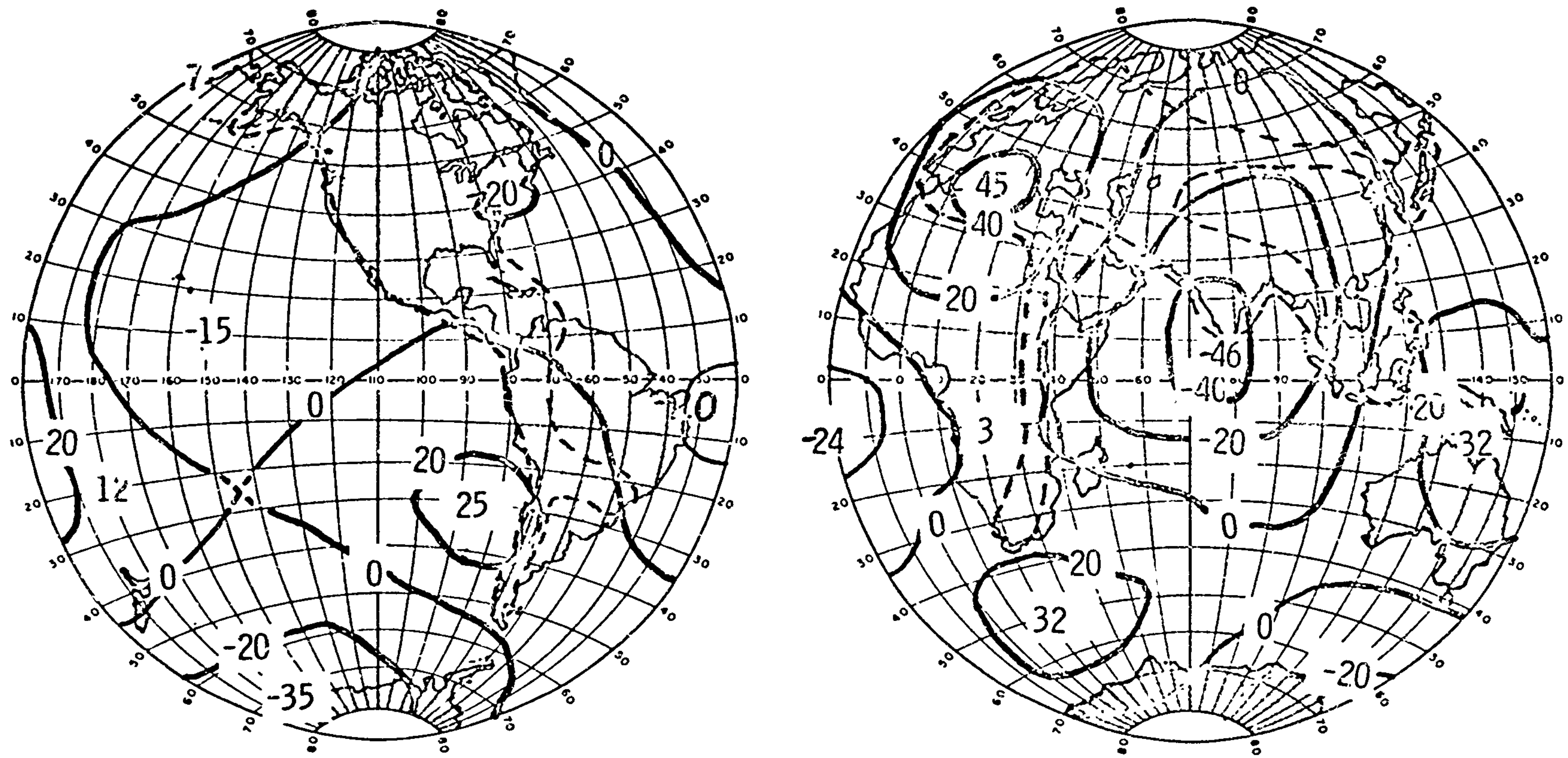


Figure 15. Geoid Heights in Meters Based on Harmonic Coefficients Determined from Terrestrial Gravimetry and Astro-Geodetic Data by Kaula (1961)

picture have gone by the board. The other main features are substantial elevations centered in France and the South Atlantic, and a depression centered near the South Pole, south of Australia. The pear shape effect is still retained, because the South Pole is depressed by some 20 meters while the North Pole is elevated by about 15 meters; but the pear has been slightly twisted, so that the maximum elevations and depressions occur away from the Poles.

These geoid maps will no doubt keep changing in detail as the years pass, but the main features are probably correct.

There are some interesting speculations which have been drawn by O'Keefe.[29] If the interior of the earth were behaving like a fluid, its flattening would be near $\frac{1}{299.8}$ rather than $\frac{1}{298.24}$; so if we wish to see how the earth departs from the fluid equilibrium, the map of its shape should be drawn not relative to the "best" ellipsoid of flattening $\frac{1}{298.24}$ but relative to an ellipsoid of flattening $\frac{1}{299.8}$. When this is done, the depression near India disappears, but an even stronger elevation than before occurs in the East Indies. This is the most active volcanic region in the world and since other volcanic regions also tend to be associated with elevated parts of the new geoid, the possibility arises that the earth's surface geometry may be linked with its vulcanology. Thus a study of space seems to be leading us to a better understanding of the interior of the earth as well.

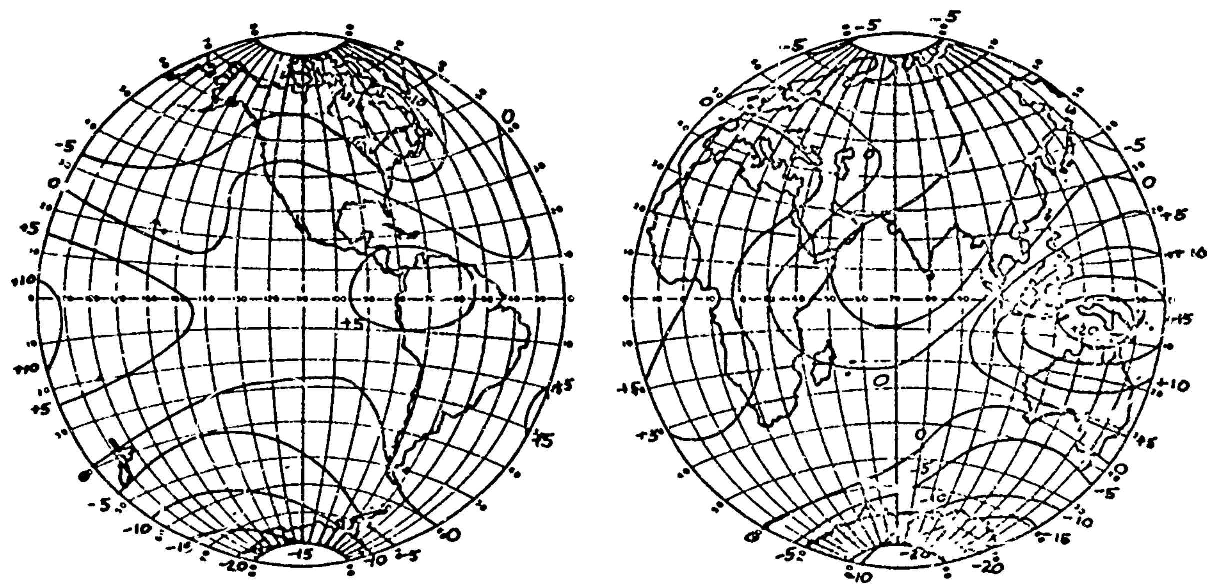


FIGURE 16—Gravity anomalies, in milligals, derived from satellite perturbations and referred to an ellipsoid with a flattening of 1/299.8.

LIST OF REFERENCES

1. Jung, K. "Figur der Erde," Handbuch der Physik, Vol XLVII, Geophysik I, Springer-Verlag, Berlin, 1956, p 543.

2. Hagihara, Y. "Recommendations on Notation of the Earth's Potential," Astrom. Journal, Vol 67, No 1, p 108, 1962.

3. Jeffreys, H. The Earth, Cambridge Univ. Press, 1959, p 136.

4. Guier, W. H. and Newton, R. R. "The Earth's Gravity Field," Rpt TG-634, Johns Hopkins Univ. Applied Physics Lab (APL), Silver Springs, Md., Dec. 1964.

5. Newton, I. Principia, Book III, Prop XIX (1687).

6. Wolaver, L. E. "Factors in Precision Space Flight Trajectories," Space Technology Lecture Series WADC TR 59-732 (1958).

7. Vignaud, H., "Histoire Critique de la Grande Entreprise de Christophe Colomb," Welter, Paris 1911, Vol I, p 311.

8. Merson. R. H. "The Perturbation of a Satellite Orbit in an Axi-symmetric Gravitational Field," RAE Tech Note No. Space 26 (1963).

9. King-Hele, D. C., G. E. Cook and J. M. Rees, "Determination of the Even Harmonics in the Earth's Gravitational Potential," Geophys Journ., Vol 8, p 119 (1963).

10. O'Keefe, J. A., A. Eckels and R. K. Squires, 1959, "The Gravitational Field of the Earth," Astron. Journ., Vol 64, pp. 245-253.

11. Todhunter, I., "A History of the Mathematical Theories of Attraction and the Figure of the Earth," Vol I, MacMillian, London, 1873.

12. Churchill, R. V. "Fourier Series and Boundary Value Problems," Chapter 9.

13. King-Hele, D. C. in "The Use of Artificial Satellites for Geodesy," edited by Vers. John Wiley, 1963, p 316.

14. Kozai, Y., p 305, ibid.

15. Newton, R. R. and Guier, W. H. "Determination of the Odd Zonal Harmonics from the Motions of Four Satellites," Johns Hopkins Univ., APL TG-519.

16. Kaula, W. M. "Determination of Variations in Gravitational Field for Calculation of Orbit Perturbations," 31 July 1964, Aerospace Corp Rpt BSD-TDR-64-104 AD 457911.

17. The Earth Today edited by A. H. Cook and T. F. Gaskell - Interscience, 1961.

18. Guier, W. H. "Determination of the Non-zonal Harmonics of the Geopotential from Satellite Doppler Data," Nature, Vol 200, No 124 (1963).

19. Michielsen, H. F. "The Odd Harmonics of the Earth's Gravitational Potential," Advances in the Astronautical Science, Vol 8, Plenum Press.

20. King-Hele, D. G., G. E. Cook and J. M. Rees, "Determination of the Even Harmonics in the Earth's Gravitational Potential," Feb 1963, RAE Tech Note No Space 25 AD 406119.

21. Kozai, Y. "Numerical Results from Orbits," Smithsonian Astrophysical Observatory Special Report 101 (1962).

22. Mercer, R. J., "Physical Constants for Satellite Calculation," Aerospace Report No. TDR-469 (5110-02)-2, Jan 1965.

23. Guier, W. H. and R. R. Newton, "The Earth's Gravity Field," Rpt No TG-634, Johns Hopkins Univ., APL Silver Springs, Md., Dec 1964.

24. Kaula, W. M. "Determination of Variation in the Gravitational Field for Calculation of Orbit Perturbations," Aerospace Corp BSD-TDR-64-104, July 1964, TDR-269(4922-10)-2.

25. Wagner, C. A. "The Drift of a 24-Hour Equatorial Satellite Due to an Earth Gravity Field through 4th Order," NASA TN-D-2103, Feb 1964.

26. Wagner, C. A. "Determination of the Ellipticity of the Earth's Equator from Observations on the Drift of the Syncom II Satellite," NASA TN D-2759, May 1965.

27. O'Keefe, J. A., Paper Presented at Conference on Artificial Satellites, Blacksburg, Virginia, Aug 1963.

28. Russell, Dugan and Stewart, "Astronomy Vol I The Solar Systems," Revised Edition, Ginn and Company, 1947.

29. Kaula, W. M., NASA TN-D-702, May 1961, AD 255523.

30. Brouwer, D. "Solution of the Problem of Artificial Satellite Theory without Drag," Astron. Joun., Vol 64, pp 378-397, 1959.

31. Cook, G. E. "The Effect of an Oblate, Rotating Atmosphere on the Orientation of a Satellite Orbit," RAE Ministry of Aviation Report, 1960.

32. Groves, G. V. "Motion of a Satellite in the Earth's Gravitational Field," Proc. Royal Soc. A, Vol 254, pp 48-65, 1960.

33. King-Hele, D. C. "The Effect of the Earth's Oblateness on the Orbit of a Near Satellite," Proc. Royal Soc. A, Vol 247, pp 49-72, 1958.

34. Kozai, Y. "On the Effects of the Sun and Moon upon the Motion of a Close Earth Satellite," Smithsonian Astrophysical Observatory Special Rpt No 22, pp 7-10, 1959.

35. Kozai, Y. "The Motion of a Close Earth Satellite," Astron. Journ., Vol 64, pp 367-377, 1959.

36. Merson, R. H. "The Motion of a Satellite in an Axi-symmetric Gravitational Field," Geophysic Journal, Vol 4, No 17, 1960.

37. Message, P. J. "On Mr King-Hele's Theory of the Effect of the Earth's Oblateness on the Orbit of a Close Satellite," Monthly Notices Royal Astron. Soc., Vol 121, pp 1-4, 1960.

38. Messen, P. "The Influence of Solar Radiation Pressure on the Motion of an Artificial Satellite," J. Geophysical Research, Vol 65, pp 1391-1396, 1960.

39. Yionoulis, S. M. "A Study of the Resonance Effects Due to the Earth's Potential Function," Journal Geophysical Research, Vol 70, No 24, Dec 1965, p 5991.

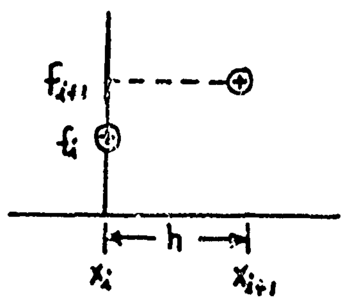
APPENDIX I

NUMERICAL INTEGRATION

Just to complete the story of special perturbation we will briefly and simply inquire into the fundamentals of numerical integration. The basic idea of integration is one of finding the area under a given curve. With differential equations this curve is not precisely known and must be approximated as we move along in time. The curve is evaluated at a few known points and then fitted to a known curve, usually a polynomial. Let's consider some very simple minded cases and then advance onward and upward.

First try to fit the points by a straight line.

Figure I-1



$$f(x) = Ax + B$$

We want to find a curve of the form Ax + B which goes through the points (o,f_i) and (h,f_{i+1}). Hence

$$B = f_i \quad ; \quad A = \frac{f_{i+1} - f_i}{h} .$$

We then proceed to integrate to obtain

$$I_1 = \int_0^h f(x)dx = \int_0^h (Ax + B)dx = A\frac{x^2}{2} + BX \Big|_0^h$$

$$I_1 = A\frac{h^2}{2} + Bh = \frac{f_{i+1} - f_i}{h}\left(\frac{h^2}{2}\right) + f_i h$$

$$I_1 = \frac{h}{2}\left(f_i + f_{i+1}\right) \qquad ; \qquad f_{i+1} = \frac{2I_1}{h} - f_i$$

This is Euler's integration formula. In terms of a differential equation we are estimating f_{i+1} by

$$f_{i+1} = f_i + h\frac{df_i}{dx}$$

given at x_i.

This can be thought of as the first few terms of a Taylor's series, i.e.,

$$f_{i+1} = f_i + \frac{h}{1!}f_i' + \frac{h^2}{2!}f_i'' + \frac{h^3}{3!}f_i''' + \cdots$$

where we have neglected the higher order terms, i.e., have truncated the series. The terms dropped are of order h^2 which we write as $O(h^2)$.

$$f_{i+1} = f_i + hf_i' + O(h^2) .$$

An error is incurred because we truncated the series. The magnitude of the error depends on h. This error is called the truncation error.

Now try fitting the curve with $Ax^2 + Bx + C$. We now require three points. We calculate A, B and C by requiring that the function

$$f(x) = Ax^2 + Bx + C$$

go through the points $(-h,f_{i-1})$, $(0,f_i)$ and (h,f_{i+1})

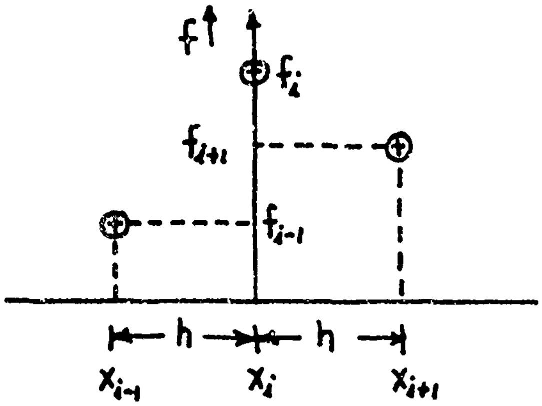


FIGURE I-2

This requires that

$$A = \frac{f_{i+1} - 2f_i + f_{i-1}}{2h^2} \; ; \quad B = \frac{f_{i+1} - f_{i-1}}{2h} \; ; \quad C = f_i \quad .$$

The integral is

$$I = \int_{-h}^{h} f(x)\,dx = \frac{2h^3}{3} A + 2hC = \frac{h}{3}\left[f_{i+1} + 4f_i + f_{i-1}\right]$$

This is Simpson's 1/3 rule for numerical integration.

To become more accurate we can try higher order polynomials. Before venturing into this area we first look at finite differences.

In developing Simpson's rule and others, one uses such quantities as

$$f_{i+1} - f_i, \qquad f_i - f_{i-1}, \qquad \text{etc.}$$

These are called finite differences. We can make a table of successive differences from any listing of a function. For example consider the function

$$f(x) = x^3 - 3x - 23$$

The differences ∇^i are obtained by subtracting each value of f from the one immediately below it ∇^1 = first difference, ∇^2 = second difference, etc.

x	f(x)	∇^1	∇^2	∇^3	∇^4
-3	-41				
		+16			
-2	-25		-12		
		+4		+6	
-1	-21		-6		0
		-2		+6	
0	-23		0		0
		-2		+6	
+1	-25		+6		0
		+4		+6	
+2	-21		-12		0
		+16		+6	
+3	-5		+18		0
		+34		+6	
+4	+29		+24		
		+58			
+5	+87				

These successive differences ∇^i are closely related to successive derivatives of f(x). The differences in general converge and they converge more rapidly as h decreases.

Now consider how we make use of these differences. First expand f_i in a Taylor's series to give

$$f_{i+1} = f_i + \frac{h}{1!} f_i' + \frac{h^2}{2!} f_i'' + \frac{h^3}{3!} f_i''' + \cdot \cdot \cdot \cdot$$

Using operator notation we may write this as

$$f_{i+1} = f_i + \frac{h}{1!} Df_i + \frac{h^2}{2!} D^2 f_i + \frac{h^3}{3!} D^3 f_i + \cdots$$

$$f_{i+1} = f_i \left(1 + \frac{h}{1!} D + \frac{h^2}{2!} D^2 + \frac{h^3}{3!} D^3 + \cdots\right) = e^{hD} f_i$$

since

$$e^{\pm x} = 1 \pm \frac{x}{1!} + \frac{x^2}{2!} \pm \frac{x^3}{3!} + \cdots$$

so that

$$e^{hD} = 1 + \frac{h}{1!} D + \frac{h^2}{2!} D^2 + \frac{h^3}{3!} D^3 + \cdots$$

So

$$f_{i+1} = e^{hD} f_i$$

likewise

$$f_{i-1} = e^{-hD} f_i \ .$$

We can then write the first differences as

$$\nabla f_i = f_i - f_{i-1} = \left[1 - e^{-hD}\right] f_i$$

so that symbolically at least, we can write

$$\nabla = 1 - e^{-hD}$$

$$e^{-hD} = 1-\nabla$$

$$\ln e^{-hD} = -hD = \ln(1 - \nabla) = -\left(\nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \cdots\right)$$

Since

$$\ln (1 \pm x) = \pm x - \frac{x^2}{2} \pm \frac{x^3}{3} - \frac{x^4}{4} \pm \frac{x^5}{5} + \cdots$$

from which we can write

$$hD = \nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \cdots\cdots$$

Now merely by multiplying the expressions out we can form the successive powers of hD as below.

$$hD = \nabla + \frac{\nabla^2}{2} + \frac{\nabla^3}{3} + \frac{\nabla^4}{4} + \cdots$$

$$h^2D^2 = \nabla^2 + \nabla^3 + \frac{11}{12}\nabla^4 + \frac{5}{6}\nabla^5 + \cdots$$

$$h^3D^3 = \nabla^3 + \frac{3}{2}\nabla^4 + \frac{7}{4}\nabla^5 + \cdots$$

$$h^4D^4 = \nabla^4 + 2\nabla^5 + \frac{17}{6}\nabla^6 + \cdots$$

$$h^5D^5 = \nabla^5 + \frac{5}{2}\nabla^6 + \frac{25}{6}\nabla^7 + \cdots$$

So - returning to Taylor's series we can write

$$f_{i+1} = f_i + h \quad [f_i' + \frac{h}{2} f_i'' + \frac{1}{6} h^2 f_i''' + \cdots]$$

$$f_{i+1} = f_i + h \quad [f_i' + \frac{hD}{2} f_i' + \frac{h^2D^2}{6} f_i' + \cdots]$$

$$f_{i+1} = f_i + h \quad [f_i' + \frac{1}{2}(\nabla + \frac{1}{2}\nabla^2 + \frac{1}{3}\nabla^3 + \frac{1}{4}\nabla^4 + \frac{1}{5}\nabla^5 + \cdots) f_i'$$

$$+ \frac{1}{6}(\nabla^2 + \nabla^3 + \frac{11}{12}\nabla^4 + \frac{5}{6}\nabla^5 + \cdots) f_i' \quad +$$ (continued next page)

$$+ \frac{1}{24} (\nabla^3 + \frac{3}{2} \nabla^4 + \frac{7}{4} \nabla^5 + \cdots) f_i'$$

$$+ \frac{1}{120} (\nabla^4 + 2\nabla^5 + \frac{17}{6} \nabla^6 + \cdots) f_i'$$

$$+ \frac{1}{720} (\nabla^5 + \frac{5}{2} \nabla^6 + \frac{25}{6} \nabla^7 + \cdots) f_i' + \cdots]$$

or --

$$f_{i+1} = f_i + h \,[1 + \frac{1}{2}\nabla + \frac{5}{12} \nabla^2 + \frac{3}{8}\nabla^3 + \frac{251}{720} \nabla^4 + \frac{95}{288} \nabla^5 + \frac{19,087}{60,480} \nabla^6$$

$$+ \cdots] f_i'$$

This is Adam's recurrence formula with f_i' at x_i given by the differential equation.

This latter formula is called the open type formula and if the n^{th} differences are used one can show the truncation error is $O(h^{n+3})$.

It is also possible to derive a closed-type formula assuming one knows f_{i+1}'. This results in

$$f_{i+1} = f_i + h\,[1 - \frac{1}{2}\nabla - \frac{1}{21}\nabla^2 - \frac{1}{24}\nabla^3 - \frac{19}{720}\nabla^4 - \frac{3}{160}\nabla^5 - \cdots]\, f'_{i+1}$$

and again one can determine the order of the neglected terms to be $O(h^{n+2})$.

The predictor-corrector technique makes use of both of these formulae. One uses the open series formula to compute f_{i+1} and this together with the differential equation gives

$$f'_{i+1} = F(x_{i+1}, f_{i+1})$$

and a new set of finite differences $\nabla^k f'_{i+1}$ is formed in order to calculate with the closed formula thereby getting a second value for f_{i+1}. This process is repeated with smaller values of h until two values of f_{i+1} do not differ by more than some preselected error criteria.

Thus the predictor-corrector technique gives an estimate for the truncation error at each stage. In addition one can allow an increase in h if the truncation error is less than this preassigned tolerance. This will reduce the total number of integration steps and thereby reduce the round-off error.

Initially we have position and velocity only at t = 0. It is therefore necessary to first use a Runge-Kutta-Gill integration formula to

generate position and velocity at t_1, t_2, t_3, t_4, $\cdots$ t_n where n is the highest difference in the Adam's formula.

The Runge-Kutta method proceeds as follows: Given $\frac{dy}{dx} = f(x,y)$ and let h represent the interval between equally spaced values of x. Then if the initial values are x_o, y_o the first increment in y is found by the following formulae.

$$k_1 = hf(x_o, y_o)$$

$$k_2 = hf(x_o + \frac{h}{2}, y_o + \frac{k_1}{2})$$

$$k_3 = hf(x_o + \frac{h}{2}, y_o + \frac{k_2}{2})$$

$$k_4 = hf(x_o + h, y_o + k_3)$$

$$\Delta y = \frac{1}{6} (k_1 + 2k_2 + 2k_3 + k_4)$$

and

$$x_1 = x_o + h \qquad y_1 = y_o + \Delta y$$

Succeeding intervals are calculated by the same relationship. Other Runge-Kutta formulae exist including a fifth order one. The one above is fourth order. For a derivation see HILDEBRAND, F.B., "Introduction to Numerical Analysis," McGraw-Hill Book Co., New York, 1956.

One can also develop an error control for the Runge-Kutta schemes. Using a Taylor series expansion, Hildebrand shows that for a fourth order Runge-Kutta method one can write

$$y_i(t, h) = y_i(t) + \frac{1}{5!} h^5 f_i^{(v)} (\theta h) \qquad 0 < \theta < 1$$

$$y_i(t, h/2) = y_i(t) + 2 \left[\frac{1}{5!} \left(\frac{h}{2}\right)^5 f_i^{(v)} \left(\bar{\theta} \frac{h}{2}\right)\right] \quad 0 < \bar{\theta} < 1$$

where $y_i(t)$ is the true value, $y_i(t, h)$ is the value computed with step size h; $y_i(t, h/2)$ is the value computed with step size h/2. The error term in the second equation is doubled since this integration involves two steps. If we assume that the fifth derivatives have the property that, for sufficiently small h,

$$f^{(v)} (\theta h) \cong f^{(v)} \left(\bar{\theta}\frac{h}{2}\right)$$

then elimination of $y_i(t)$ from the above two equations gives the error as

$$\varepsilon_i = \frac{1}{15} [y_i(t, h) - y_i(t, \tfrac{h}{2})].$$

This gives a good estimate of the truncation error and is used as an error control criteria to reduce or increase the size of the interval h.

In addition to truncation error, one also has a round-off error. This error increases as h decreases.

The error incurred in rounding the result of each computation is propagated through all subsequent computation. When integration

is involved the errors generated at each step are propagated to the end of the integration. An actual round-off error in one integrand value will appear in all subsequent single integrals. That is, due to round-off error in the last calculation, the new interval has the wrong initial conditions for position and velocity. Some systems are extremely sensitive to initial condition errors and in such systems any error, due to truncation or round-off, will propagate rapidly.

Thus besides considering the truncation error at each step it is extremely important to know how a single truncation error propagates as the solution proceeds. This is strictly a function of the particular differential equation being solved and unfortunately is seldom if ever investigated.

Consider the simple linear equation

$$\frac{d^4y}{dt^4} + 4y = 0$$

with initial conditions of $y'''(0) = +2$, $y''(0) = -2$, $y'(0) = +1$ and $y(0) = 0$. It is easily verified that the answer is $y = e^{-t}\sin t$.

Now if one attempts to obtain a numerical solution for this differential equation, by any scheme what-so-ever, in less than 10 seconds the numerical solution does not have even one good significant figure.

See table below.

t	$y = e^{-t}\sin t$	y Numerical	Difference	% Error
0	0	0	0	0
1	0.30956	0.30956	0	0
2	0.12306	0.12306	0	0
4	-0.013861	-0.013862	.000001	-
6	-0.00069260	-0.00069146	-.00000114	-0.2%
8	+0.00033189	+0.00033841	+.00000652	+2.0%
10	-0.000024699	-0.00012732	+.000102621	+415%

And it gets worse as time goes by.

MORAL - Never feed a computer a problem whose answer you don't know in advance. It will give you the wrong answer every time.

The propagation of round-off error can sometimes be estimated. Since we sometimes round-up and sometimes round-down, one cannot consider the rounding errors to propagate directly. For want of nothing better one can assume rounding errors are Gaussian distributed about zero. With this assumption, which is perhaps a little harsh, one can show that the mean error in a single integration, due to round-off, varies as the square root of the number of steps. On double integration, the mean error goes as the three-halves power of the number of steps.

REF: D. Brouwer, "On the Accumulation of Errors in Numerical Integration," Astronomical Journal, Vol. 46, p. 149, 1937.

Thus for 10,000 integration steps, $\sqrt{10^4} = 10^2$ or the last two significant figures are most probably in error after 10,000 single integration. The bound is actually a little high but is often of the proper order of magnitude.

To show these effects, Baker has applied numerical integration to the motion about the unstable libration point in the earth-moon restricted three-body problem. If there were precisely zero error the particle should remain at this point. As error propagates, ever so slightly, the particle will leave like a herd of turtles. Figure I-3 shows this effect for various error controls. As the integration interval is decreased, time at libration point is increased until we reduce to 10^{-6}. Further reduction to 10^{-7} causes round-off error to become larger than truncation error. We are caught between the two types of errors.

REF: "Efficient Precision Orbit Computation Techniques by R. M. L. Baker, Jr., et. al., ARS Reprint No. 869-59. Presented at ARS Meeting June 8-11, 1959.

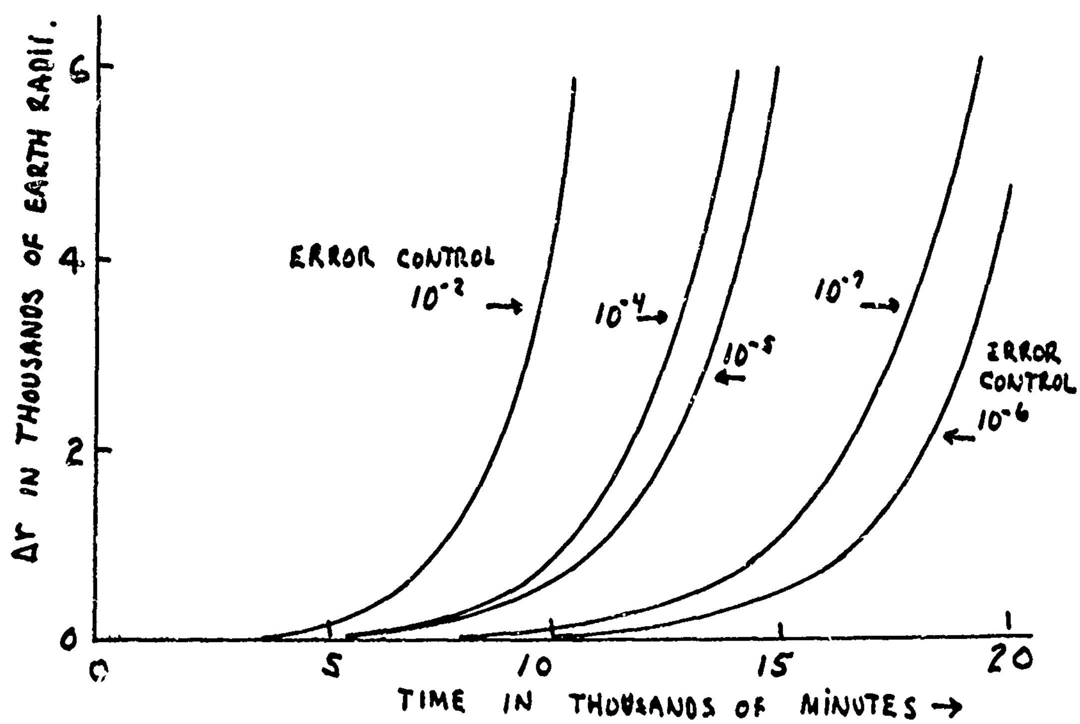


FIG. I-3 ANALYTICAL SOLUTION VS COWELL'S METHOD

Sterne develops a method adopted from Rademacher which makes use of the adjoint system to estimate the propagation of truncation and round-off error. This technique was applied by Robert Zani in an AFIT Thesis to problems in celestial mechanics.

Besides these normal integration schemes one can also use power series expansion methods. A discussion of this technique as applied to problems in dynamics is given in the NASA report by Fehlberg. He also gives higher order implicit and explicit Runge-Kutta formulae.

REFERENCES

[1] Baker, R. M. L., Jr. and Maud Makemson, "Introduction to Astrodynamics".

[2] Brouwer, Dirk and Gerald M. Clemence, "Methods of Celestial Mechanics", Academic Press, 1961.

[3] Fehlberg, Erwin, "New One-Step Integration Methods of High Order Accuracy Applied to Some Problems in Celestial Mechanics", NASA TR R-248, October 1966.

[4] Sterne, Theodore E., "An Introduction to Celestial Mechanics", Interscience Publishers, 1960.

[5] Zani, Robert C., "A Computer Study of the Estimated Propagation of Errors in the Numerical Integration of Ordinary Differential Equations", AFIT Master's Thesis GE/EE/67A-13, March 1967.

MIMIC

In the course of solving a differential equation by numerical methods, it becomes necessary to translate the mathematics into digital language such as FORTRAN, ALGOL, etc. To be most efficient the student should learn one of these languages such as FORTRAN IV. There is, however, a simple language which is limited in scope but is more than adequate for most problems in celestial mechanics or for ordinary differential equations in general. Furthermore, it is simple enough to be self taught in the course of a day. The first of these is called MIDAS and an improved version is called MIMIC. Both of these languages were developed by F. J. Sansom and H. L. Peterson of the System Engineering Group here at Wright Field.

MIDAS is a digital computer program that was designed to solve systems of ordinary differential equations. It does this by employing a block-oriented input language, that may be most easily thought of as a large Tinkertoy, a set of building blocks which the programmer puts together to conform to the particular system he wishes to simulate. Each block is capable of performing just one distinct operation, like integration, summation, multiplication, limiting, etc. Writing a MIDAS program consists of drawing a block diagram that shows the interconnection of these blocks. The interconnections are then listed, and together with a few additional instructions are keypunched and submitted for solution.

To illustrate the method of formulating a MIDAS solution, consider the following simple example. Assume that we wish to solve the second-order differential equation

$$Mx'' + Bx' + Kx = 0; \quad x'(0) = 0, \; x(0) = 20$$

for the three cases

Case 1 $M = 10$, $B = 2.5$, $K = 8.6$

Case 2 $M = 10$, $B = 3.2$, $K = 8.6$

Case 3 $M = 10$, $B = 2.5$, $K = 15.0$

and that we are interested in observing x'', x' and x for $0 \leq t \leq 5$. A block diagram that represents this system is shown in Figure I-5.

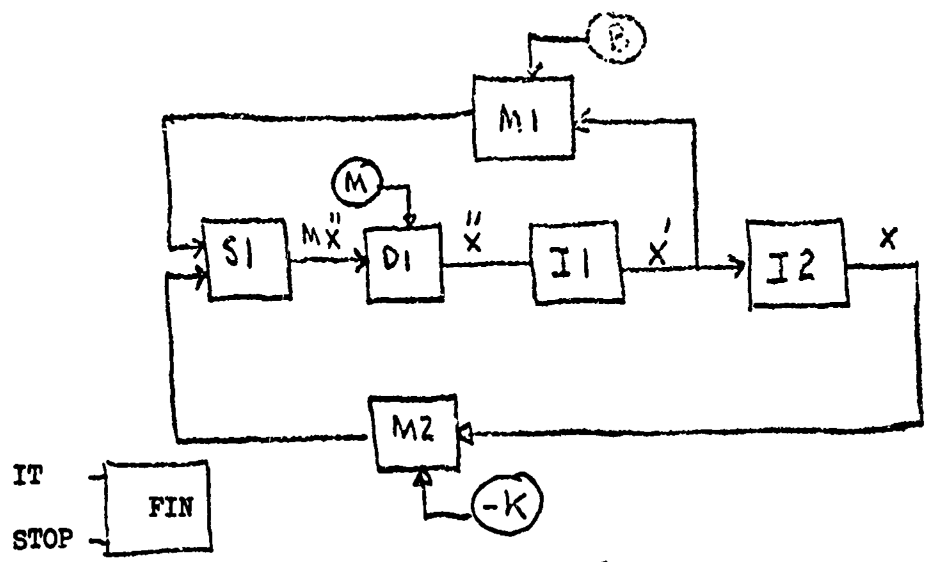


FIGURE I-5

Each block in this diagram has been named in a manner indicative of the operation it is to perform. The first letter indicates the operation and the succeeding, completely arbitrary, alphanumeric characters are used to distinguish it from other elements of the same

type. Thus I2 stands for an integrator named 2 -- it could equally well have been named IX.

The operation of S1 is summation; its output is equal to the sum of its inputs. Thus $-(B\dot{x} + Kx) = M\ddot{x}$, which is, of course, the equation we wish to solve. The rest of the diagram is concerned with operating on $M\ddot{x}$ to provide the two inputs, $-B\dot{x}$ and $-Kx$, necessary to specify S1. Hence Mx is divided by M in D1, $\ddot{x}$ is integrated in I1 to form $\dot{x}$, etc. The FIN block is a finish condition; computation on a case is stopped and the next case started when IT (the independent variable) exceeds a constant named STOP. As many finish conditions as required may be imposed.

The listing for this diagram is

```
S1     M1,M2
D1     S1,M
I1     D1
I2     I1
M1     -B,I1
M2     I2,-K
FIN    IT, STOP
```

Each line of the listing corresponds to one element in the diagram. First the name of the element, identifying it by type and number, is given; then the sources of its inputs are entered. Multiple inputs are separated by commas. This listing is not enough to solve the

problem; additional cards are necessary to call the MIDAS subroutine, enter data, and control the printing of results. Figure I-8 shows a complete listing for this program. The constants M and STOP are

```
$IBFTC  CALL
        CALL MIDAS
        END
$DATA
SOLUTION OF MASS, SPRING, DAMPER SYSTEM
        CON    M,STOP
        PAR    -B,-K
        IC     I2
        S1     M1,M2
        D1     S1,M
        I1     D1
        I2     I1
        M1     -B,I1
        M2     I2,-K
        FIN    IT,STOP
        HDR    TIME,ACC.,VEL.,DISP.
        HDR
        RO     IT,D1,I1,I2
        END
```

(continued on next page)

10.   5.

-2.5   -8.6

20.

-3.2   -8.6

20.

-2.5   -15.

20.

$EOF

FIGURE II-8

named on a constant card (CON), the parameters -B and -K on a parameter card (PAR), and the initial condition of integrator 2 on an initial condition card (IC). Printout of results is controlled by the readout (RO) statement and, unless otherwise specified, a printout occurs at every .1 unit of the independent variable. Variables named in RO statements will be printed in a fixed 6-column format. Headings for these columns can be introduced by using a header (HDR) statement. The END card is used to distinguish between MIDAS statements and data. Comments may be used at will before the END card. Any statement with a non-blank column 1 is considered to be a comment.

This simple example uses a very limited set of the type and number of elements available in MIDAS. The program is large enough to solve complex systems of nonlinear differential equations of up to order 100. A complete description of MIDAS is contained in Air Force Technical Documentary Report No. SEG-TDR064-1.[1]

MIMIC

The MIMIC program is an outgrowth of MIDAS. It attempts to eliminate some of the shortcomings of MIDAS and yet remain a simple, easily learned tool for solving parallel systems.

MIDAS requires that a block diagram, representing the system to be studied, be drawn. From this diagram a listing is prepared and the solution obtained. MIMIC obviates the necessity of the block diagram by allowing the user to name the variables in any manner that he chooses, independent of the operation used to evaluate the variable. Further, the arithmetic operators +, -, *, and /, denoting addition, subtraction, multiplication, and division respectively, have been added.

The MIMIC coding for the example given above is shown in Figure I-7. Data is named and entered, and output is handled as in MIDAS.

```
        CON(M)
        PAR(B,K)
2DX     -(B*1DX+K/X)/M
1DX     INT(2DX,0.)
X       INT(1DX,20.)
        FIN(T,5.)
        HDR(T,ACC,VEL,DISP)
        OUT(T,2DX,1DX,X)
        END
```

FIGURE I-7

The differences are fairly evident. $\ddot{x}$ is named 2DX rather than D1, $\dot{x}$ is 1DX, etc. More than one operation can be performed on one line. Constants can be identified and assigned a value by using a literal name, that is 5. is the name of a constant that has a value of 5.

In general a MIMIC program consists of a series of statements of the form

RESULT = EXPRESSION

where RESULT is a user supplied name identifying the variable evaluated by EXPRESSION. The equality sign is implicit and need not be present. EXPRESSION is a sequence of variables (RESULTS) and FUNCTIONS separated by arithmetic operators, commas, and parentheses. All FUNCTIONS are represented by three letter mnemonic codes followed by up to six arguments enclosed in parentheses and separated by commas. The arguments of FUNCTIONS may themselves be EXPRESSIONS.

As an example, consider the following equation:

$$x = y + (z + e^{-at})^{1/2}$$

It can be programmed as

```
X = Y + SQR(Z + EXP(-A*T))
```

where SQR is the square root function and EXP the exponential function. The argument of SQR is an EXPRESSION.

MIMIC provides functions for all the operations (blocks) in MIDAS, including functions for addition (ADD), multiplication (MPY), negation (NEG), and division (DIV). These last four functions are provided so that block-oriented programs can be written. Additionally, MIMIC provides logical functions, hybrid functions, and the ability to write subprograms in MIMIC language. Control over the execution of the program is extended by allowing limited branching. A LOGICAL CONTROL VARIABLE can be defined that will control the execution of individual statements. A more detailed description of MIMIC is given in SESCA Internal Memo 65-12.[2] (More readily available as Reference 3).

REFERENCES

1. R. T. Harnett, F. J. Sansom, and L. M. Warshawsky, "MIDAS Programming Guide," Report No. SEG-TDR-64-1, Systems Engineering Group, Wright-Patterson Air Force Base, Ohio, January 1964.
2. H. E. Petersen and F. J. Sansom, "MIMIC - A Digital Simulator Program," SESCA Internal Memo 65-12, Systems Engineering Group, Wright-Patterson Air Force Base, Ohio, May 1965.
3. F. J. SANSOM and H.E. PETERSON, "MIMIC PROGRAMMING MANUAL", TECHNICAL REPORT SEG-TR-67-31, JULY 1967 (AD-656 301).

APPENDIX II

Review of the General Concepts of Dynamics

Source: Douglas Aircraft Report SM-23401 by P. A. Lagerstrom and M. E. Graham, "Some General Concepts of Dynamics and Their Application to the Restricted Three Body Problem," December 1958.

"Classical Mechanics" by Herbert Goldstein Addison-Wesley, Reading, Mass., 1959.

It is assumed that the student has had a good course in the fundamentals of dynamics. This appendix, which is drawn directly from the Douglas report, is intended to be a review of conservation theorems, Lagrange's equations and an introduction to Hamilton's equations.

Dynamics may be defined as the theory of motion of matter under the influence of forces. Here we are concerned with the special branch of dynamics in which one makes the idealizing assumption that matter consists of a finite number of rigid bodies; this is opposed to dynamics of continuous systems (hydrodynamics, theory of elasticity, etc.). A further idealizing assumption is that the mass of each body is concentrated at a point. The general problem of the motion of mass points is the following: consider a system of $\underline{n}$ mass points. The i^{th} mass point has mass $\underline{m_i}$, position $\underline{\vec{r}_i}$ (radius vector) and is acted on by the force $\underline{\vec{f}_i}$. According to classical (pre-relativistic) mechanics, the motion of the particles then obeys Newton's law

$$m_i \frac{d^2\vec{r}_i}{dt^2} = \vec{f}_i \qquad i = 1, 2, \ldots n \qquad \text{(II-1)}$$

To make this system of equations complete, one needs physical laws which define the forces. The force $\underline{\vec{f}_i}$ will, in general, depend on the position and velocities of all the particles. If one considers rigid bodies of finite extension instead of mass points, one can show that equation (II-1) is still valid for the center of mass of the bodies. In addition, one then also needs a set of equations, similar in structure to equation (II-1) which govern the rotation of the bodies.

The mathematical formulation of the dynamics of rigid bodies thus leads to a system of ordinary differential equations. Continuous mechanics on the other hand leads to partial differential equations, such as the Navier-Stokes equations. In the advanced theory of rigid bodies (Hamilton-Jacobi theory, see Goldstein P 273 ff) one may associate certain partial differential equations with a system of the type (II-1). However, basically, the physical laws are formulated with the aid of ordinary differential equations. Once the explicit expressions for the $\underline{\vec{f}_i}$ are given, the physical problem is reduced to a well-defined mathematical problem. Dynamics thus becomes mathematics rather than physics. However, most force-laws lead to nonlinear expressions for the $\underline{\vec{f}_i}$ and, as a result, present-day mathematics is inadequate for the solution of many important problems. In most problems, solvable or not, it is advisable to rely heavily on physical and geometrical intuition.

If the $\underline{\vec{f}_i}$ are gravitational forces only, then the problem of solving Eq. II-1 becomes a problem in celestial mechanics. Here we are concerned only with pre-relativistic mechanics; hence, the gravitational forces are given by Newton's formula. [The special theory of relativity provides a

correction to Newton's law of motion (II-1); the general theory of relatively also replaces Newton's law of gravitation by more accurate laws.] Then, if the number of mass points is n, the problem of solving Eq. II-1 is called the n-body problem. For $n > 2$ this problem has been solved only partially. Much of the mathematical information about the n-body problem is contained in the so-called conservation laws (integrals of motion). For $n = 2$ the conservation laws provide the complete solution.

Conservation laws play an important role in many problems of dynamics other than those of celestial mechanics. In this section, their use will therefore be illustrated with the aid of simple problems. Furthermore, the conservation laws are intimately connected with the use of generalized coordinates and with Lagrange's and Hamilton's equations of motion. These concepts will also be discussed in this section in connection with simple examples.

1. Conservation of Energy -- Generalized Coordinates - Illustrated by Simple Examples

Example 1 - One-dimensional oscillator:

Here we have one particle of mass m, position x, acted upon by a force f which depends only upon the position, $f = f(x)$. [The simplest case would be a linear spring-mass system, a harmonic oscillator, for which $f(x) = -kx$, k = spring constant.] The equation of motion is

$$m \frac{d^2x}{dt^2} = f(x) \tag{II-2}$$

This second-order equation can be written as a pair of first-order equations if we define $v = dx/dt$ = velocity.

$$\frac{dx}{dt} = v \qquad \text{(II-3a)}$$

$$m\,\frac{dv}{dt} = f(x) \qquad \text{(II-3b)}$$

The function $\underline{f}(\underline{x})$ defines a force-field in (one-dimensional) space. If the particle happens to be at $\underline{x}$, then the force acting upon it is the value of the force-field at $\underline{x}$. Since only one particle is involved, the force-field is given a priori as a function of space; in the n-body problem, on the other hand, the force-field acting on one body would depend on the position of the other bodies.

We define a function $\underline{V}(\underline{x})$ by

$$V(x) = -\int_0^x f(x)\,dx \qquad \text{(II-4)}$$

$\underline{V}(\underline{x})$ will be called the potential energy, or the force potential, for reasons to be given later. A potential can always be defined from a one-dimensional force field; in higher dimensions, it can be defined only when the force field is irrotational.* [Note that for the harmonic oscillator $\underline{V}(\underline{x}) = \underline{kx}^2/2$.]

The choice of the lower limit in Eq. II-4, equivalent to specifying a constant of integration, is unimportant; in each problem one picks a convenient value. The essential part of the definition (II-4) is the differential law

$$-\frac{dV}{dx} = f(x) \qquad \text{(II-5)}$$

---

*The force field $\vec{f}(\vec{r})$ is irrotational if $\oint \vec{f} \cdot d\vec{r} = 0$ for any closed path.

If the particle is at $\underline{x} = \underline{x_o}$ when $\underline{t} = \underline{t_o}$ it is said to have the potential energy $\underline{V}(\underline{x_o})$. The potential energy of a moving particle is then an implicit function of time. The time rate of change of the potential energy of a moving particle is then given by the substantial derivative

$$\frac{dV}{dt} = \lim_{\Delta t \to o} \frac{V[x(t + \Delta t)] - V[x(t)]}{\Delta t} \tag{II-6}$$

The rules of differential calculus then give

$$\frac{dV}{dt} = \frac{dV}{dx}\frac{dx}{dt} = -fv \tag{II-7}$$

The following calculation shows the importance of the function $\underline{V}(\underline{x})$: If both sides of Eq. II-3b are multiplied by $\underline{v}$ one finds

$$mv\frac{dv}{dt} = fv \tag{a}$$

Hence, using Eq. II-7,

$$\frac{d}{dt}\left(\frac{mv^2}{2}\right) + \frac{dV}{dt} = 0 \tag{II-8}$$

or

$$T + V = \text{constant} = E \tag{II-9}$$

where

$$T = \frac{mv^2}{2} = \text{kinetic energy}$$

Equation II-9 is a first example of a conservation law. It states that the total energy ($\underline{T} + \underline{V}$) is conserved; i.e., unchanged, during the motion of the particle. (The fact that $\underline{f}$ depended on $\underline{x}$ only, and not on time, was used implicitly above; time dependent force fields and potential functions will be considered later.)

The total energy is also called an integral of the motion. The reason for the name is that in deriving the conservation law (II-9) we have integrated the equations of motion once. The original second-order system (II-3a, b) has been replaced by one first-order equation involving one constant of integration. This is seen by rewriting II-9 as

$$\frac{dx}{dt} = \sqrt{\frac{2}{m}[E - V(x)]} \tag{II-10}$$

Since $\underline{V}$ is not explicitly a function of $\underline{t}$, Eq. II-10 can be immediately solved in the sense that $\underline{t}$ may be found as a function of $\underline{x}$,

$$t = \int_0^x \frac{dx}{\sqrt{\frac{2}{m}[E - V(x)]}} + t_o \tag{II-11}$$

where $\underline{t_o}$ is a second constant of integration.

Thus a general solution of the system (II-3a, b) involving two constants of integration, $\underline{E}$ and $\underline{t_o}$, has been obtained. If at time $\underline{t} = 0$, the position $\underline{x} = \underline{x_o}$ and velocity $\underline{v} = \underline{v_o}$ are given, the constants $\underline{E}$ and $\underline{t_o}$ are

$$E = \frac{mv_o^2}{2} + V(x_o), \qquad t_o = -\int_0^{x_o} \frac{dx}{\sqrt{\frac{2}{m}[E - V(x)]}} \tag{II-12}$$

In principle, Eq. II-11 may, of course, be inverted to give $\underline{x}$ as a function of $\underline{t}$.

We will carry through the integration for the special case of the harmonic oscillator for which

$$f(x) = -kx, \qquad V(x) = kx^2/2 \tag{b}$$

Then

$$t-t_o = \int_o^x \frac{dx}{\sqrt{\frac{2}{m}\left(E - \frac{kx^2}{2}\right)}} = \frac{\sqrt{\frac{2E}{k}}}{\sqrt{\frac{2E}{m}}} \int_o^x \frac{d\left(\sqrt{\frac{k}{2E}}\,x\right)}{\sqrt{1 - \frac{kx^2}{2E}}} \tag{c}$$

or, with $\omega^2 = \underline{k}/\underline{m}$,

$$t-t_o = \frac{1}{\omega} \text{ arc sin}\left(\sqrt{\frac{k}{2E}}\,x\right) \tag{d}$$

or

$$x = \sqrt{\frac{2E}{k}} \sin \omega(t-t_o) \tag{II-13}$$

If we consider the special initial conditions

$$x = x_o, \qquad v = v_o = o \quad \text{for} \quad t = o \tag{e}$$

then

$$E = kx_o^2/2, \qquad t_o = -\frac{\pi}{2}/\omega, \qquad \text{and} \qquad x = x_o \cos \omega t \; . \tag{f}$$

For the case of a harmonic oscillator, the final results are, of course, easily obtained directly from the second-order equation of motion. However, if the restoring force is non-linear, the method of solution shown above is very useful.

Example 2 - Two-dimensional roller coaster:

This example will follow the same method of solution (but adapted to two dimensions) as was used in Example 1, and will introduce the concepts of constrained motion and of generalized coordinates. A heavy particle moves without friction, but is constrained by tracks to follow the surface of a roller coaster defined by $y = G(x)$. The forces acting are a uniform gravitational force (in the $-y$ direction) and the force of constraint, which is normal to the path. The position and velocity of the particle are given by the vectors $\underline{\vec{r}}$ and $\underline{\vec{v}}$:

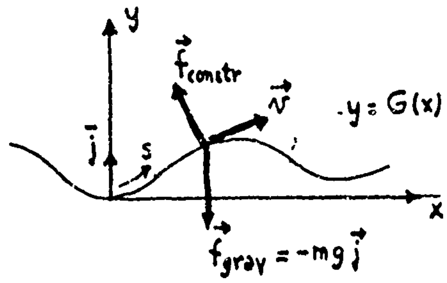


Fig. II-1: Particle moving on roller coaster

$$\vec{r} = x\vec{i} + y\vec{j}, \qquad \vec{v} = \frac{d\vec{r}}{dt} = \frac{dx}{dt}\vec{i} + \frac{dy}{dt}\vec{j} \tag{II-14}$$

The equations of motion are (cf. Eqs. II-3a,b):

$$\frac{d\vec{r}}{dt} = \vec{v} \tag{II-15a}$$

$$m\frac{d\vec{v}}{dt} = \vec{f}_{grav} + \vec{f}_{constr} \tag{II-15b}$$

The force of constraint, $\underline{\vec{f}}_{constr}$, is unknown. However, without knowing $\underline{\vec{f}}_{constr}$, the method of Example 1 again leads to a first integral of the motion (a conservation-of-energy theorem).

First, as in Example 1, we will define a potential $\underline{V}$ at each point of space. Note that in Example 1, $\underline{V}$, as defined in Eq. II-4, can be interpreted as the negative of the work done by the force on the particle in moving the particle from 0 to x. Now in the present example the force of constraint acts normal to the particle path so that it does no work; only the gravitational force works. One then defines the potential $\underline{V}$ by

$$-\text{grad } V = \vec{f}_{grav} \tag{II-16a}$$

or

$$-\vec{i}\,\frac{\partial V}{\partial x} - \vec{j}\,\frac{\partial V}{\partial y} = -mg\vec{j} \tag{II-16b}$$

Thus, choosing a convenient constant of integration, we may write

$$V = mgy = mgG(x) \tag{II-17}$$

Now apply the same trick as was used in Example 1; i.e., take the dot (scalar) product of each side of Eq. II-15b with $\vec{\underline{v}}$. (The vector equation then becomes a scalar equation.)

$$m\,\frac{d\vec{v}}{dt}\cdot\vec{v} = \vec{f}_{grav}\cdot\vec{v} + \vec{f}_{constr}\cdot\vec{v} \tag{a}$$

or $$\frac{d}{dt}\left(\frac{m}{2}\,\vec{v}\cdot\vec{v}\right) = -mg\vec{j}\cdot\left[\frac{dx}{dt}\,\vec{i} + \frac{dy}{dt}\,\vec{j}\right] + \vec{f}_{constr}\cdot\vec{v} \tag{b}$$

Noting that $\vec{v}\cdot\vec{v} = v^2$ and $\vec{f}_{constr}\cdot\vec{v} = 0$ since $\vec{f}_{constr}$ is normal to $\vec{v}$, one gets

$$\frac{d}{dt}\left(\frac{mv^2}{2} + mgy\right) = 0 \tag{II-18}$$

or $\quad T + V = E$ (II-19)

with $T = mv^2/2$ = kinetic energy, $V = mgy$ = potential energy, and E = constant. Thus again we have a theorem of energy conservation.

We will now solve this problem by reducing it to the same form as the problem of the non-linear oscillator of Example 1. The conservation law (II-8) of Example 1 can be written

$$\frac{m}{2}\frac{d}{dt}\left[\left(\frac{dx}{dt}\right)^2\right] + \frac{dV(x)}{dt} = 0 \tag{c}$$

while the conservation law (II-18) of the present example can be written

$$\frac{m}{2}\frac{d}{dt}\left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2\right] + \frac{dV(y)}{dt} = 0 \tag{d}$$

In order to transform (d) into an equation formally identical with (c), we introduce the element of path length ds:

$$(ds)^2 = (dx)^2 + (dy)^2 \tag{II-20}$$

Then the length s along the curve $y = G(x)$ is

$$s = \int_o^s ds = \int \sqrt{(dx)^2 + (dy)^2}$$

$$= \int_o^x \sqrt{1 + \left(\frac{dy}{dx}\right)^2}\, dx = \int_o^x \sqrt{1 + \left(\frac{dG(x)}{dx}\right)^2}\, dx \tag{II-21}$$

Thus the path length s has been defined as a function of x, or implicitly x has been defined as a function of s, so that all the quantities of (d)

may be expressed as functions of s. The conservation law for the roller coaster then becomes

$$\frac{m}{2}\left(\frac{ds}{dt}\right)^2 + \frac{dV(s)}{dt} = 0 \qquad \text{(II-22a)}$$

where V(s) is the potential energy expressed as a function of s,

$$V(s) = mgG[x(s)] \qquad \text{(II-22b)}$$

Equation II-22 is formally identical with Eq. II-8. Thus there is an exact mathematical analogy between the two-dimensional constrained frictionless motion on the roller coaster and the one-dimensional undamped motion of an oscillator (linear or non-linear). As an example of this analogy, we shall find the roller coaster corresponding to a linear spring. For the linear spring, $V(x) = kx^2/2$. Hence, for the corresponding roller coaster y as a function of s should be given by

$$mgy = ks^2/2 \qquad \text{(a)}$$

To find y as a function of x, we rewrite (a) as

$$mgy = \frac{k}{2}\left(\int_0^x \sqrt{1+\left(\frac{dy}{dx}\right)^2}\, dx\right)^2 \qquad \text{(b)}$$

Taking the square root and differentiating, one finds the differential equation,

$$\sqrt{\frac{2a}{y}}\ \frac{dy}{dx} = \sqrt{1+\left(\frac{dy}{dx}\right)^2} \qquad \text{(c)}$$

where $\quad a = \frac{mg}{4k}$ (d)

or $\quad \frac{dx}{dy} = -\sqrt{\frac{2a - y}{y}}$ (e)

This formula has an indeterminacy in the sign of the square root. To remove this indeterminacy, we will use a parametric representation. An angle ϕ is introduced by

$$-\tan \frac{\phi}{2} = \sqrt{\frac{2a - y}{y}} \qquad \text{(f)}$$

Then $\frac{dx}{dy} = -\tan \frac{\phi}{2}$ (g)

and the shape of the roller coaster can be defined parametrically by

$$y = 2a \cos^2 \frac{\phi}{2} \qquad \text{(h)}$$

$$x = \int_{\pi}^{\phi} \left(-\tan \frac{\phi}{2}\right) \left(-2a \cos \frac{\phi}{2} \sin \frac{\phi}{2} \, d\phi\right) \qquad \text{(i)}$$

or $\quad 2 = a(\phi - \pi - \sin \phi)$ (II-23a)

$\quad y = a(1 + \cos \phi)$ (II-23b)

Equation II-23 is the parametric representation of the cycloid shown in Fig. II-2. A wheel of radius a rolls without slipping on the underside of the dashed line $y = 2a$. A point P on this wheel then describes the cycloid given by Eq. II-23 if P is chosen as shown in the figure. The parameter ϕ is the turning angle normalized so that $\phi = 0$ for $x = -\pi a$.

Note that the period of a linear oscillator is amplitude-independent. It then follows that the period of a particle moving on a roller coaster

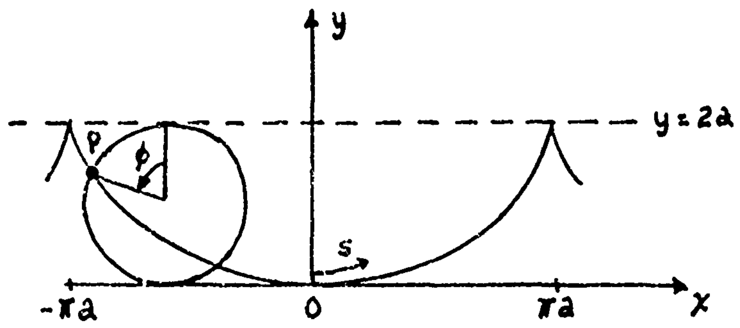


Fig. II-2: Cycloid

shaped like a cycloid is independent of amplitude (except that the amplitude, of course, is limited by the geometry of the cycloid). Since the period of the harmonic oscillator is $T = 2\pi\sqrt{m/k}$, it follows from (d) that the period $\underline{T}$ of a particle moving on the cycloid described by Eq. II-23 or Fig. II-2 is

$$T = 4\pi\sqrt{\frac{a}{g}} \tag{j}$$

This amplitude independence was discovered by Christian Huygens (published in 1673) in connection with a watch-making project.

Note that in the original formulation (II-15) of the roller coaster problem, two spatial coordinates were used. However, because of the constraint there is only one degree of freedom and only one coordinate should be necessary. It was actually found that the length-parameter $\underline{s}$ was a convenient coordinate. This is an example of a "generalized" coordinate. The Lagrangian and Hamiltonian formulations of the basic

equations of motion are especially suitable for handling problems in which generalized coordinates are used. This will be described under subheadings 2. to 5. of this section.

Example 3 - Rigid pendulum in plane:

Consider the rigid pendulum as shown in Fig. II-3. A point of mass m is suspended by a weightless string. The extension of the string due to tension is neglected. In the case of planar motion, the mass point is restricted to motion on a circle of radius ℓ. This is then actually a special case of motion on a roller coaster. Thus, the conservation law (II-19) applies:

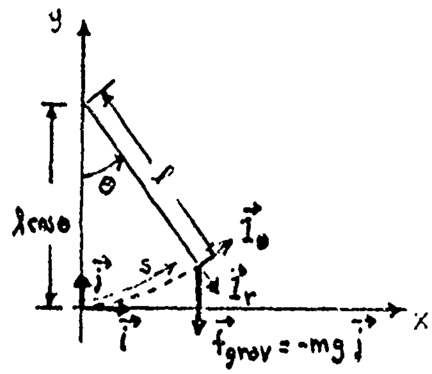


Fig. II-3: Pendulum

$$T + V = E \qquad \text{(II-19)}$$

or

$$\frac{m}{2}\left(\frac{ds}{dt}\right)^2 + mgy = E$$

In the notation of Fig. II-3, one finds: path length $s = \ell\theta$, velocity $v = \ell d\theta/dt$, $y = \ell(1 - \cos\theta)$. Then, in terms of the generalized coordinate* θ, Eq. II-19 becomes

*Note that a generalized coordinate need not have the dimension of length.

$$\frac{m}{2}\,\ell^2 \left(\frac{d\theta}{dt}\right)^2 + mg\ell\,(1 - \cos\theta) = E \qquad \text{(II-24)}$$

For θ small, $1 - \cos\theta \approx \theta^2/2$ and Eq. II-24 describes a harmonic oscillation.

Equation II-24 can, of course, be derived directly from Newton's equations in the same manner as was used to derive the conservation laws (II-9) and (II-19) of Examples 1 and 2. Newton's equation, with the acceleration in polar coordinates given by Eq. I-10, is

$$\vec{f}_{grav} + \vec{f}_{constr} = -\frac{m}{\ell}\left(\frac{\ell d\theta}{dt}\right)^2 \vec{1}_r + m\ell\,\frac{d^2\theta}{dt^2}\,\vec{1}_\theta \qquad \text{(a)}$$

The radial component of (a) is useful only if one wishes to know the tension in the string. The tangential component is

$$-mg\,\sin\theta = m\ell\,\frac{d^2\theta}{dt^2} \qquad \text{(II-25)}$$

Multiplying both sides by $\ell d\theta/dt$,

$$-mg\ell\,\sin\theta\,\frac{d\theta}{dt} = \frac{m}{2}\,\ell^2\,\frac{d}{dt}\left(\frac{d\theta}{dt}\right)^2 \qquad \text{(b)}$$

$$\frac{d}{dt}\left[\frac{m}{2}\,\ell^2\left(\frac{d\theta}{dt}\right)^2 + mg\ell\,(1 - \cos\theta)\right] = 0 \qquad \text{(c)}$$

Equation (c) is obviously equivalent to Eq. II-24.

2. Examples of Lagrange's Formulation of the Equations of Motion

Lagrange's equations of motion are a mathematical reformulation of Newton's basic physical law (II-1). These equations are especially convenient for the case of constrained motion and for many other cases in which generalized coordinates are used. They lead in a natural way to Hamilton's equations (under subheading 4). Lagrange's equations will be

introduced with the aid of some simple examples. The conditions under which the equations are valid will be discussed later, but a general derivation of the equations will not be given. (See, e.g., Goldstein, p. 14 ff.)

The Lagrangian $\underline{L}$ is defined to be

$$L = T - V \tag{II-26}$$

Note that $\underline{L}$ is the kinetic energy minus the potential energy. Lagrange's equation of motion with one degree of freedom is

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0 \tag{II-27}$$

where $\underline{L}$ is considered a function of $\underline{x}$ and $\underline{\dot{x}}$ ($\dot{x} \equiv dx/dt$) so that $\partial L/\partial \dot{x}$ is obtained holding $\underline{x}$ fixed, and vice versa.

It will be shown, for examples 1 and 3 of Subsection 1, that Lagrange's equations are equivalent to Newton's equations of motion. The Lagrange equations will also be written for more complicated examples, but without showing their equivalence with Newton's equations.

Example 1 - One dimensional oscillator - Lagrangian formulation:

$$\text{Lagrangian:} \quad L = m\,\frac{\dot{x}^2}{2} - V(x) \tag{II-28}$$

Then, $\partial L/\partial \dot{x} = m\dot{x}$, $\partial L/\partial x = -\ \partial V/\partial x$.

$$\text{Lagrange's equation:} \ \frac{d}{dt}\,(m\dot{x}) + \frac{\partial V}{\partial x} = 0 \tag{II-29}$$

or $\quad m \dfrac{d^2x}{dt^2} - f(x) = 0$

Thus, here, Lagrange's equation is equivalent to Newton's equation of motion (II-3).

Example 3 - Rigid pendulum in plane - Lagrangian formulation:

This is again a problem with one degree of freedom. The Lagrangian is a function of the generalized coordinate θ and its time derivate $\dot{\theta}$.

$$\text{Lagrangian:} \quad L \equiv T - V = \frac{m\ell^2}{2} \dot{\theta}^2 - mg\ell\,(1 - \cos\theta) \tag{II-30}$$

Then $\partial L/\partial\dot{\theta} = m\ell^2\dot{\theta}$, $\partial L/\partial\theta = -mg\ell \sin\theta$.

$$\text{Lagrange's equation:} \quad \frac{d}{dt}\left(\frac{\partial L}{\partial\dot{\theta}}\right) - \frac{\partial L}{\partial\theta} = 0 \tag{II-31}$$

or $\quad \dfrac{d}{dt}\,(m\ell^2\dot{\theta}) + mg\ell \sin\theta = 0$

Lagrange's equation is again equivalent to the corresponding Newton's equation (II-25).

* * *

In each of the two previous examples, the motion was allowed only one degree of freedom, and the system was conservative. Knowing T and V, it would have been sufficient to write (T + V) = constant to describe the motion. Thus, Lagrange's equations in no way simplified the problems. However, if there is more than one degree of freedom, then more than one equation is needed to determine the motion. One could, of course, add

the necessary number of Newton's equations. However, this may not be convenient if they involve the unknown forces of constraint. On the other hand, if one chooses suitable non-cartesian coordinates so that the constraints don't appear, then the acceleration components may not be easy to write. In the following examples, it appears simpler to use the Lagrangian formulation of the equations of motion.

<u>Example 4</u> - Arbitrary motion in space (described in spherical coordinates) by Lagrangian formulation:

The cartesian coordinates (x, y, z) are related to the spherical coordinates (r, ϕ, θ) of Fig. II-4 by the transformation

$$x = r \sin\phi \cos\theta$$
$$y = r \sin\phi \sin\theta \qquad \text{(II-32)}$$
$$z = r \cos\phi$$

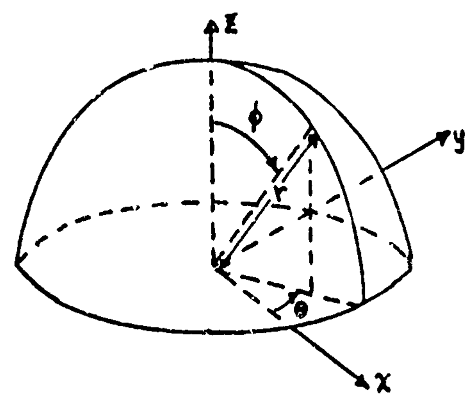


Fig. II-4

(The angles ϕ and θ correspond to the co-latitude and longitude on the earth's surface.) For motion under no constraints, the generalized coordinates are r, ϕ, θ so that the Lagrangian L is a function of six variables, $r, \phi, \theta, \dot{r}, \dot{\phi}, \dot{\theta}$, and three Lagrange equations are required to describe the motion.

The kinetic energy is

$$T = \frac{m}{2}\left[\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2\right] = \frac{m}{2} v^2 \qquad \text{(a)}$$

$$= \frac{m}{2}\left[\dot{r}^2 + r^2 \dot{\phi}^2 + r^2 \sin^2\phi \, \dot{\theta}^2\right]$$

[T as a function of $\dot{r}$, $\dot{\phi}$, $\dot{\theta}$ (given above) can be found by taking time derivatives of the coordinate transformation (II-32). However, it is found more easily by noting that $v^2 = (ds/dt)^2$ where the length element ds has the components, in the r, ϕ, θ directions, of dr, r$d\phi$, r sin ϕ $d\theta$.]

We assume that the forces are derivable from a time-independent potential V,

$$\vec{f} = -\text{grad } V \tag{b}$$

or in components

$$f_r = -\frac{\partial V}{\partial r}, \qquad f_\phi = -\frac{1}{r}\frac{\partial V}{\partial \phi}, \qquad f_\theta = -\frac{1}{r \sin\phi}\frac{\partial V}{\partial \theta} \tag{c}$$

The partial derivatives of the Lagrangian (L = T-V) are then

$$\begin{aligned}
\partial L/\partial\dot{r} &= m\dot{r}, \; \partial L/\partial r = mr\dot{\phi}^2 + mr \sin^2\phi\, \dot{\theta}^2 + f_r \\
\partial L/\partial\dot{\phi} &= mr^2\dot{\phi}, \; \partial L/\partial\phi = mr^2 \sin\phi \cos\phi\, \dot{\theta}^2 + rf_\phi \\
\partial L/\partial\dot{\theta} &= mr^2 \sin^2\phi\, \dot{\theta}, \; \partial L/\partial\theta = r \sin\phi\, f_\theta
\end{aligned} \tag{d}$$

Then the Lagrange equations are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) - \frac{\partial L}{\partial r} \equiv m\left[\frac{d}{dt}\dot{r} - r\dot{\phi}^2 - r\sin^2\phi\, \dot{\theta}^2\right] - f_r = 0 \tag{II-33a}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\phi}}\right) - \frac{\partial L}{\partial \phi} \equiv mr\left[\frac{\frac{d}{dt}(r^2\dot{\phi}) - r^2 \sin\phi \cos\phi\, \dot{\theta}^2}{r}\right] - r\, f_\phi = 0 \tag{II-33b}$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) - \frac{\partial L}{\partial \theta} \equiv mr \sin\phi\left[\frac{\frac{d}{dt}(r^2 \sin^2\phi\, \dot{\theta})}{r \sin\phi}\right] - r \sin\phi\, f_\theta = 0 \tag{II-33c}$$

The three terms in brackets must be the three components of the accelerations, since these equations should be equivalent to Newton's equations of motion. These acceleration components could be computed by the method used to compute the acceleration in polar coordinates which depends upon computing the changes in the base vectors: $\partial \vec{1}_r/\partial \theta$, etc. However, the method used above requires only that one knows how to express the length element ds in the r, ϕ, θ coordinates. Thus, in this example, the Lagrangian formulation provides a shortcut to finding the accelerations.

<u>Example 3</u> - Rigid pendulum in plane - special case of Example 4 with $\phi = \pi/2$:

The generalized coordinates are now r, θ which are really polar coordinates in the plane z = 0. The motion is described by Eqs. II-33a, c of Example 4 with $\phi = \pi/2$. The terms in brackets are the acceleration components, and it can be checked that they agree with the acceleration derived by using base vector derivatives.

<u>Example 5</u> - Spherical pendulum (rigid pendulum in space) by Lagrangian formulation:

This is a special case of Example 4 in which the constraint r = constant is introduced. There are, then, only two degrees of freedom, and the generalized coordinates are ϕ, θ. The motion is described by Eqs. II-33b, c of Example 4 (with <u>r</u> constant). The force of constraint, if desired, can be found from Eq. II-33a (with <u>r</u> constant).

* * *

As can be seen in the above examples, the Lagrangian formulation of the equations of motion was useful when the motion was expressed in non-cartesian coordinates and when the motion was constrained. The use of a Lagrange equation for the one-dimensional oscillator problem may, of course, seem unnecessarily artificial; however, for the spherical pendulum, e.g., it was a considerable simplification to write the equations of motion in Lagrange's form.

3. General Discussion of Lagrange's Equations - Further Examples

In Examples 1 and 3 we wrote the Lagrange equations of motion, and then checked that they were equivalent to Newton's equations; but in Examples 4 and 5 we accepted Lagrange's equations without such a check. We shall now give the general form of Lagrange's equations and discuss conditions for their validity.

First we define a holonomic constraint: If there are n particles, the first with cartesian coordinates x_1, x_2, x_3, the second with coordinates x_4, x_5, x_6, and etc., then the particles move under holonomic constraints if the conditions of constraint may be expressed by m equations

$$f_j\,(x_1, x_2, x_3; x_4, x_5, x_6; \cdots\ x_{3n-2}, x_{3n-1}, x_{3n}, t) = 0 \qquad \text{(II-34)}$$

where $j = 1, 2, \cdots m$; $m<3n$

Thus, a holonomic constraint can be expressed as a definite equation involving only the coordinates of the particles and perhaps time. (For example, the constraint on the motion of a rigid spherical pendulum is

holonomic, and can be expressed as $x_1^2 + x_2^2 + x_3^2 - \ell^2 = 0$. For a planar pendulum, there is the additional holonomic constraint, $x_3 = 0$.) A non-holonomic constraint might depend on the velocities of the particles, or be expressible only as an inequality. An example of a non-holonomic constraint is a sphere rolling without slipping on a flat surface. (Goldstein, p. 12 ff.)

If a system of n particles is moving under m holonomic constraints, then only 3n-m coordinates are required to describe the motion; i.e., the system has $\ell = 3n-m$ degrees of freedom. Then, with the aid of the m equations of the type (II-34), ℓ generalized coordinates $q_1, q_2, \ldots q_\ell$ may be introduced by means of the transformation equations.

$$x_k = x_k\,(q_1, q_2, \cdots q_\ell, t) \tag{II-35}$$

where $k = 1, 2, \cdots 3\,n$

The existence of this set of transformation equations (II-35) is a condition for the validity of the Lagrange equations.

Before stating the consequences of the existence of the transformation equations (II-35), we digress to discuss generalized coordinates and generalized forces. The concept of generalized coordinates as introduced by Eq. II-35 seems very abstruse. Actually the generalized coordinates may be any independent set of coordinates which are adequate to describe the motion. They are distinguished from the cartesian coordinates in that they are not connected by any constraints. (If there were no constraints, the cartesian coordinates would be a perfectly good set of

generalized coordinates. However, even in the case of no constraints, the symmetry properties of a problem may make it advisable to introduce non-cartesian coordinates, say, e.g., spherical coordinates.) They also have the characteristic that they need not have the dimensions of length. We have already met generalized coordinates in some of the examples of Subsection 2: motion on the roller coaster (Example 2) was described in terms of $\underline{s}$, the length along the track; the motion of the plane pendulum (Example 3) in terms of $\underline{\theta}$; and the motion of the spherical pendulum (Example 5) in terms of the co-latitude and longitude, $\underline{\phi}$ and $\underline{\theta}$.

To each generalized coordinate q_j there corresponds a generalized force component Q_j which will now be defined. First consider the work δW that would be done by the actual forces if they were to displace the particles by infinitesimal amounts δx_k where the δx_k satisfy the constraints (II-34) but are otherwise arbitrary. Actually each δx_k is a possible displacement of one cartesian component of one particle. (These δx_k are called "virtual displacements" and are to be distinguished from actual displacements which occur in time. That is, a virtual displacement provides a comparison between two positions x_k and $x_k + \delta x_k$ at the same time.) The work δW (called "virtual work") is then

$$\delta W = \sum_{k=1}^{3n} f_k \, \delta x_k \qquad \text{(II-36)}$$

The force f_k is one cartesian component of the total force acting upon one particle. It may include a force of constraint (which, in general, is unknown, a priori), but the constraining force need not be included: by definition, a constraining force on any particle acts in a direction

in which the particle is not free to move in a virtual displacement.* Thus a constraining force does no work in a virtual displacement.**

The virtual work (II-36), with the aid of Eq. II-35, can be written in terms of generalized coordinates:

$$\delta W = \sum_{k=1}^{3n} f_k \left[\sum_{j=1}^{\ell} \frac{\partial x_k}{\partial q_j} \delta q_j \right] \tag{a}$$

or, regrouping the terms,

$$\delta W = \sum_{j=1}^{\ell} \delta q_j \sum_{k=1}^{3n} f_k \frac{\partial x_k}{\partial q_j} \tag{b}$$

Since δW is independent of the constraining forces, one can now define a generalized force Q_j which is independent of the constraining forces:

$$Q_j = \sum_{k=1}^{3n} f_k \frac{\partial x_k}{\partial q_j} \tag{II-37}$$

The virtual work is then

$$\delta W = \sum_{j=1}^{\ell} Q_j \, \delta q_j \tag{II-38}$$

(Note that Q_j need not have the dimensions of a force, but that $Q_j \, \delta q_j$ always has the dimensions of work.)

---

*Note that, by this definition, sliding friction forces, e.g., would not qualify as constraining forces since they may do work in a virtual displacement. However, the derivation of Eq. II-40 is perfectly valid for the case with friction if the friction is included in the f_k.

**If the equation of constraint (II-34) contains time explicitly, then the force of constraint might do work in an actual displacement although it does no work in a virtual displacement.

To illustrate the concept of a generalized force as defined by Eq. II-37, we return to two of the examples discussed previously.

Example 3 - Rigid pendulum in plane - generalized force:

Let f_x and f_y be the cartesian components of the total force acting on the pendulum. The generalized coordinate is the angle θ and the transformation equations are (see Fig. II-3)

$$x = \ell \sin \theta, \qquad y = \ell (1 - \cos \theta) \tag{a}$$

The actual forces are gravity and the constraint (tension in string):

$$\vec{f}_{grav} = - mg\vec{j}, \quad \vec{f}_{constr} = (-f_{constr} \sin \theta)\vec{i} + (f_{constr} \cos \theta)\vec{j} \tag{b}$$

Then the generalized force Q_θ, as defined by Eq. II-37, is

$$Q_\theta = f_x \frac{dx}{d\theta} + f_y \frac{dy}{d\theta}$$

$$= (-f_{constr} \sin \theta) (\ell \cos \theta) + (f_{constr} \cos \theta - mg) (\ell \sin \theta)$$

$$\text{or} \qquad Q_\theta = - mg\ell \sin \theta \tag{II-38}$$

Note that the constraining force drops out of the equation and Q_θ can be identified as the moment of force, or the torque, on the pendulum (ℓ · tangential force of Eq. II-25).

Example 2 - Two-dimensional roller coaster - generalized force:

The generalized coordinate for this case is the path length s (see Fig. II-1). The cartesian components of the actual forces -- gravity and the constraint -- are

$$f_x = -f_{constr} \frac{dy}{ds}, \qquad f_y = f_{constr} \frac{dx}{ds} - mg \tag{a}$$

Then the generalized force Q_s is by definition (II-37)

$$Q_s = f_x \frac{dx}{ds} + f_y \frac{dy}{ds}$$

$$= \left(-f_{constr} \frac{dy}{ds}\right) \frac{dx}{ds} + \left(f_{constr} \frac{dx}{ds} - mg\right) \frac{dy}{ds}$$

$$\text{or} \qquad Q_s = -mg \frac{dy}{ds} \tag{II-39}$$

Again, the unknown force of constraint has dropped out. For this example, Q_s can be identified as the tangential component of the actual force.

* * *

We now return to the general case of a system of n particles with ℓ degrees of freedom. We describe the system by ℓ generalized coordinates q_j. Corresponding to these generalized coordinates are ℓ generalized forces Q_j (which do not involve the unknown forces of constraint). The next step is to find ℓ differential equations relating the Q_j to the q_j. These equations of motion are derived in Goldstein . We will here merely state: if the transformation equations (II-35) exist, then Newton's 3n equations of motion may be replaced by the following set of ℓ equations.

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_j}\right) - \left(\frac{\partial T}{\partial q_j}\right) = Q_j, \qquad j = 1, 2, \cdots \ell \tag{II-40}$$

where T is the kinetic energy of the system and Q_j is defined by Eq. II-37.

Forces dependent only upon coordinates and time, and derivable from a potential function:

In Eq. II-40, the generalized forces Q_j were obtained from actual forces f_k which had not been restricted in any way. Now let us restrict the actual forces (not including forces due to constraints which by definition act normal to any virtual displacements) to those which are expressible as the gradient of a potential $\underline{V}$ where $\underline{V}$ depends only on the positions of the particles and perhaps on time. (At this point, we are thus excluding such forces as sliding friction which would depend on the velocities.) That is, we assume the f_k are given by equations of the form (II-41):

$$f_k\,(x_1,\ x_2,\ \cdots\ x_{3n},\ t) = -\frac{\partial V}{\partial x_k}\,(x_1,\ x_2,\ \cdots\ x_{3n},\ t) \qquad \text{(II-41)}$$

Equation II-41 can be written in terms of generalized forces and coordinates by multiplying both sides by $\partial x_k/\partial q_j$ and summing over k:

$$Q_j \equiv \sum_{k=1}^{3n} f_k \frac{\partial x_k}{\partial q_j} = -\sum_{k=1}^{3n} \frac{\partial V}{\partial x_k} \frac{\partial x_k}{\partial q_j} \qquad \text{(a)}$$

or

$$Q_j = -\frac{\partial V}{\partial q_j}\,(q_1,\ q_2,\ \cdots\ q_\ell,\ t) \qquad \text{(II-42)}$$

We illustrate Eq. II-42 with the plane pendulum and roller coaster (Examples 3 and 2). For both cases the potential function $V = mgy$. For the plane pendulum, $y = \ell\,(1 - \cos\theta)$ where θ is the generalized coordinate. Then Q_θ as defined by Eq. II-42 is

$$Q_\theta = -\frac{d}{d\theta}\,[mg\ell\,(1 - \cos\theta)] = -mg\ell\,\sin\theta \tag{II-38}$$

For the roller coaster, the generalized coordinate is s; and Q_s, as defined by Eq. II-42, is

$$Q_s = -\frac{d}{ds}\,(mgy) = -mg\,\frac{dy}{ds} \tag{II-39}$$

These obviously agree with the results obtained from definition (II-37).

If the generalized forces are of the form (II-42), then the equations of motion (II-40) can be written in terms of T and V. Since V is velocity-independent, Eq. II-40 can be written,

$$\frac{d}{dt}\left[\frac{\partial}{\partial \dot{q}_j}\,(T - V)\right] - \frac{\partial T}{\partial q_j} = -\frac{\partial V}{\partial q_j} \tag{b}$$

Introducing the Lagrangian L = T-V one obtains Lagrange's equations:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right) - \frac{\partial L}{\partial q_j} = 0 \tag{II-43}$$

where $L = L(q_1, q_2 \cdots q_\ell, \dot{q}_1, \cdots \dot{q}_\ell, t)$.

Example 6 - One-dimensional oscillator with time-dependent force:

In the discussion above we considered the case for which the Lagrangian may depend on time. A simple case of this will now be discussed in detail. We consider the one-dimensional oscillator (Example 1) but make now the assumption that the force field f depends on time t as well as on space x. As in Eqs. II-4 we define a potential

$$V(x, t) = - \int_{o}^{x} f(x, t)dx \tag{II-44a}$$

or

$$- \frac{\partial V}{\partial x} = f(x, t) \tag{II-44b}$$

Let the particle have positions x(t) and x(t + Δt) at two infinitesimally close instants. The difference in the value of V for the particle at these two instants is

$$\Delta V = V[x(t + \Delta t), t + \Delta t] - V[x(t), t]$$

$$= \frac{\partial V}{\partial x} \Delta x + \frac{\partial V}{\partial t} \Delta t \tag{a}$$

Hence the substantial derivative of V is

$$\frac{dV}{dt} = \frac{\partial V}{\partial x}\frac{dx}{dt} + \frac{\partial V}{\partial t} = -fv + \frac{\partial V}{\partial t} \tag{b}$$

Multiplying the equation of motion

$$m \frac{dv}{dt} = f \tag{c}$$

by v one finds

$$\frac{d\left[\frac{mv^2}{2}\right]}{dt} - fv = 0 \tag{d}$$

or

$$\frac{d}{dt} (T + V) = \frac{\partial V}{\partial t} \tag{II-45a}$$

Hence the total energy is not conserved; its time-rate of change is $\partial V/\partial t$.

Lagrange's equations of motion are, however, still valid. Defining

$$L = T - V = \frac{mv^2}{2} - V(x, t) = L(x, v, t) \qquad \text{(e)}$$

one finds

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v}\right) - \frac{\partial L}{\partial x} =$$

$$\frac{d}{dt}(mv) + \frac{\partial V}{\partial x} = m\frac{dv}{dt} - f = 0 \qquad \text{(c)}$$

Note that Eq. II-45a may be written

$$\frac{d}{dt}(T + V) = -\frac{\partial L}{\partial t} \qquad \text{(II-45b)}$$

Example 7 - Bead sliding on rotating wire - time-dependent constraint - time dependent force

We assume that a rigid wire rotates with constant angular velocity $\underline{\omega}$ and that a bead slides on this wire without friction. As generalized coordinate we use $\underline{r}$, the distance from the origin. The equations connecting the cartesian coordinates $\underline{x}$, $\underline{y}$, and the generalized coordinate $\underline{r}$ are, for a suitable choice of the time origin

$$x = r\cos\omega t, \qquad y = r\sin\omega t \qquad \text{(II-46)}$$

These equations are a special case of Eqs. II-35. Note that in the present case $\underline{x}$ and $\underline{y}$ depend on the generalized coordinate and on time. It is assumed that (x, y) is an inertial system so that only real forces need be considered. The frictional force is neglected.

First we assume that there are no forces except the force of constraint. Then

$$T = \frac{m}{2}\left[\dot{r}^2 + (r\omega)^2\right] \qquad \text{(II-47a)}$$

$$V = 0 \qquad \text{(II-47b)}$$

and hence

$$\frac{\partial T}{\partial \dot{r}} = m\dot{r}, \qquad \frac{\partial T}{\partial r} = mr\omega^2 \qquad \text{(II-48)}$$

Since L = T, Lagrange's equations state that

$$m\left(\frac{d^2r}{dt^2} - r\omega^2\right) = 0 \qquad \text{(II-49)}$$

Eq. II-49 states that the radial acceleration is zero which is Newton's law of motion for the case of zero radial force.

Secondly, we make the assumption that the force of gravity acts on the bead. Then Eq. II-47b is replaced by

$$V = mgy = mgr \sin \omega t \qquad \text{(II-50)}$$

Hence the generalized force Q_r is

$$Q_r = -\frac{\partial V}{\partial r} = -mg \sin \omega t \qquad \text{(II-51)}$$

and Lagrange's equation states that

$$m\left(\frac{d^2r}{dt^2} - r\omega^2\right) + mg \sin \omega t = 0 \qquad \text{(II-52)}$$

This is again Newton's law of motion. The generalized force Q_r is actually the radial component of the force of gravity.

<u>Generalization to include</u> velocity-dependent forces:

In the previous derivation of Lagrange's equation (II-43), it was assumed that the force-field was derivable from a potential $V(q_1, \ldots q_\ell, t)$ with the generalized force given by

$$Q_j = -\frac{\partial V}{\partial q_j} \qquad \text{(II-42)}$$

Thus velocity-dependent forces such as an electro-magnetic Lorentz force were excluded. It is, however, obvious from the derivation that if one replaces Eq. II-42 by the more general equation

$$Q_j = \frac{d}{dt}\left(\frac{\partial V}{\partial \dot{q}_j}\right) - \frac{\partial V}{\partial q_j} \qquad \text{(II-53)}$$

then Lagrange's equations are still valid. In Eq. II-53, it is assumed that $\underline{V}$ may depend on $q_1, \ldots q_\ell, \dot{q}_1, \ldots \dot{q}_\ell$, and $\underline{t}$. If $\underline{V}$ does not depend on the velocities $\dot{q}_i$ then Eq. II-53 reduces to Eq. II-42.

We have thus obtained a mathematically trivial extension of the validity of Lagrange's equations. However, it so happens that this extension has physical significance. This can be seen from the discussion of the following example.

<u>Example 8</u> - Charged particle in a given electro-magnetic field - velocity-dependent force: (Goldstein, pp. 19-21)

Let the given electro-magnetic field be characterized by $\vec{E}$, the electric field strength, and $\vec{B}$, the magnetic induction. From Maxwell's equations for a vacuum, one may show that $\vec{E}$ and $\vec{B}$ may be derived from a scalar potential ϕ and a vector potential $\vec{A}$ by

$$\vec{E} = -\text{grad}\ \phi - \frac{1}{c}\frac{\partial \vec{A}}{\partial t}, \qquad c = \text{speed of light} \tag{II-54a}$$

$$\vec{B} = \text{curl}\ \vec{A} \tag{II-54b}$$

If a particle of charge q is moving in this field, the force exerted by the field on the particle is

$$\vec{f} = \vec{f}_{el} + \vec{f}_{magn} \tag{II-55a}$$

where

$$\vec{f}_{el} = q\vec{E} \qquad = \text{electrostatic force} \tag{II-55b}$$

$$\vec{f}_{magn} = \frac{q}{c}(\vec{v} \times \vec{B}) \qquad = \text{Lorentz force} \tag{II-55c}$$

$\vec{v}$ = velocity of the particle

If one puts

$$V = q\left(\phi - \frac{1}{c}\vec{A} \cdot \vec{v}\right) \tag{II-56}$$

one may show that $\vec{f}$ actually is given by Eq. II-53, i.e.,

$$f_x = \frac{d}{dt}\left(\frac{\partial V}{\partial v_x}\right) - \frac{\partial V}{\partial x}, \text{ etc.} \qquad \text{(II-57}$$

4. Hamilton's Equations of Motion - Examples

Hamilton's equations of motion have some features in common with Lagrange's equations. They are mathematical reformulations of Newton's law of motion which are convenient to use when constraints are present or when, for one reason or another, one introduces coordinates other than the original cartesian coordinates of the individual mass points. They have a certain advantage over Lagrange's equations when one is deriving and discussing conservation laws. As stated above, in classical dynamics, the use of Hamilton's equations is a matter of mathematical technique; no new physical principles are introduced. As a digression, it may be pointed out, however, that in the transition from classical dynamics to quantum mechanics, one uses the Hamiltonian formulation of the former as a starting point

Hamilton's equations are easily derived from Lagrange's equations. Before carrying out this derivation we shall, however, illustrate their use with the aid of several examples.

<u>Example 1</u> - One-dimensional oscillator - Hamiltonian formulation: (p3)

The one-dimensional oscillator was discussed in Example 1. The Lagrangian was

$$L = \frac{m\dot{x}^2}{2} - V(x) \tag{II-28}$$

We now define

$$\text{Generalized momentum} = p_x = \frac{\partial L}{\partial \dot{x}} \tag{II-58}$$

$$\text{Hamiltonian} = H = p_x \dot{x} - L \tag{II-59}$$

The Lagrangian is a function of $\underline{x}$ and $\dot{x}$. One may use Eq. II-58 to express $\underline{\dot{x}}$ as a function of $\underline{x}$ and p_x. Hence $\underline{H}$ may be considered as a function of $\underline{x}$ and p_x. We now claim that the equations of motion are

$$\frac{dx}{dt} = \frac{\partial H}{\partial p_x}, \qquad \frac{dp_x}{dt} = -\frac{\partial H}{\partial x} \tag{II-60}$$

These are Hamilton's equations. Furthermore, we claim that Eqs. II-60 imply that the Hamiltonian is a constant of motion

$$\frac{dH}{dt} = 0 \qquad \text{or} \qquad H = \text{constant} \tag{II-61}$$

The above statements will now be verified. From Eqs. II-28 and II-58, one finds

$$p_x = m\dot{x} \tag{a}$$

Thus p_x is the ordinary momentum; for the present example the adjective "generalized" is superfluous. Furthermore, from Eqs. II-59, II-28 and (a),

$$H = p_x \, \dot{x} - \frac{m\dot{x}^2}{2} + V \tag{b}$$

or

$$H = \frac{p_x^2}{2m} + V = T + V \tag{c}$$

In the present example H is thus the total energy. Equations II-60 are then

$$\frac{dx}{dt} = \frac{p_x}{m} \tag{d}$$

$$\frac{dp_x}{dt} = - \frac{\partial V}{\partial x} = f \tag{e}$$

These are indeed Eqs. II-3a, b written in a slightly different notation.

Since H is equal to T + V we know already (Eq. II-9) that it is a constant of the motion. However, it is of importance to derive this formally from Eqs. II-60 without reference to their physical meaning. The derivative in Eq. II-61 is the substantial time-derivative. Its meaning is familiar to aerodynamicists; it will, however, be restated here. If a particle moves, then at each instant of time it has a definite position $\underline{x}$ and a definite momentum p_x. Hence, for a given motion, $\underline{x}$ and p_x are functions of time. Since $\underline{H}$ is completely determined by $\underline{x}$ and p_x only, it may thus be considered as depending on time only for a given motion. The increase in H from time = t to time = t + Δt for a given motion is then, to first order in Δt

$$\begin{aligned} \Delta H &= H[x(t + \Delta t),\ p_x(t + \Delta t)] - H[x(t),\ p_x(t)] \\ &= \frac{\partial H}{\partial x} \Delta x + \frac{\partial H}{\partial p_x} \Delta p_x = \frac{\partial H}{\partial x} \frac{dx}{dt} \Delta t + \frac{\partial H}{\partial p_x} \frac{dp_x}{dt} \Delta t \end{aligned} \tag{f}$$

Hence

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial p_x}\frac{dp_x}{dt} \tag{g}$$

Inserting into this equation the values of dx/dt and dp_x/dt as given by Hamilton's equations (II-60), one derives Eq. II-61.

Hamilton's equations have an obvious hydrodynamical analogy. Consider a fluid particle in the (x, p)-plane. Its velocity components are then dx/dt and dp/dt. Equations II-60 then state that these velocities may be derived from a "stream function" H. The statement dH/dt = 0 means in this interpretation that the value of the stream function is constant along a streamline.

<u>Example 3</u> - Rigid pendulum in a plane - Hamiltonian formulation: The Lagrangian for this problem was found to be

$$L = \frac{m\ell^2\dot{\theta}^2}{2} - mg\ell\,(1 - \cos\theta) \qquad \text{(II-30)}$$

As before we define

$$\text{Generalized momentum} = p_\theta = \frac{\partial L}{\partial \dot{\theta}} \qquad \text{(II-62)}$$

$$\text{Hamiltonian} = H = p_\theta\,\dot{\theta} - L \qquad \text{(II-63)}$$

and claim validity for Hamilton's equations,

$$\frac{d\theta}{dt} = \frac{\partial H}{\partial p_\theta}, \qquad \frac{dp_\theta}{dt} = -\frac{\partial H}{\partial \theta} \qquad \text{(II-64)}$$

Furthermore, we claim that <u>H</u> is a constant of the motion.

The generalized momentum is in the present case

$$p_\theta = m\ell^2\dot{\theta} \qquad \text{(a)}$$

We see that the term "generalized" is necessary in the present example. The actual momentum is $m\ell\dot{\theta}$. Hence, p_θ is the moment of momentum about the point of suspension, i.e., the angular momentum. (Cf. this with

the statement that the generalized force corresponding to the coordinate $\underline{\theta}$ is the moment of the force, i.e., the torque.) The Hamiltonian is

$$
\begin{aligned}
H &= \frac{m\ell^2\dot{\theta}^2}{2} + mg\ell\,(1 - \cos\theta) \\
&= \frac{p_\theta^2}{2m\ell^2} + mg\ell\,(1 - \cos\theta) = T + V
\end{aligned}
\tag{b}
$$

Hamilton's equations of motion are then

$$\frac{d\theta}{dt} = \frac{p_\theta}{m\ell^2} \tag{c}$$

$$\frac{dp_\theta}{dt} = -mg\ell \sin\theta \tag{d}$$

The first equation defines p_θ in terms of the angular velocity $\dot{\theta} = (d\theta/dt)$. The second equation is then identical with the previously derived equation

$$m\,\frac{d(\ell\dot{\theta})}{dt} = -mg\sin\theta \tag{II-25}$$

which states that the tangential acceleration is matched by the tangential component of the gravitational force. Finally, we observe that Eqs. II-60 are formally identical with Eqs. II-64, the symbols $\underline{x}$ and p_x in the former corresponding to the symbols $\underline{\theta}$ and p_θ in the latter. One then proves $dH/dt = 0$ exactly as before.

<u>Example 6</u> - One-dimensional oscillator with time-dependent force -- Hamiltonian formulation:

Hamilton's equations for the oscillator were given above (Example 1) under the assumption that the force depends on $\underline{x}$ only. We now assume

that $\underline{f}$ depends on $\underline{t}$ as well as on $\underline{x}$. It is easily checked that Hamilton's equations (II-60) are still valid. Furthermore H = T + V. However, the conservation law (II-61) is no longer valid. Since now $L = L(x, \dot{x}, t)$, it follows that $H = H(x, p_x, t)$. From differential calculus, it then follows that

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\frac{dx}{dt} + \frac{\partial H}{\partial p_x}\frac{dp_x}{dt} + \frac{\partial H}{\partial t} \tag{a}$$

Note that $\partial H/\partial t$ is computed by keeping $\underline{x}$ and p_x fixed and varying time; the substantial derivative dH/dt is computed by following the particle, i.e., by letting the variation of $\underline{x}$ and p_x be determined by the motion of the particle. These derivatives are thus conceptually different. However, Hamilton's equations (II-60) show that their value is the same, i.e.,

$$dH/dt = \partial H/\partial t \tag{II-65}$$

From the definition of H (II-59) it follows that

$$\partial H/\partial t = - \partial L/\partial t \tag{II-66}$$

Hence

$$\frac{d}{dt}(T + V) = - \frac{\partial L}{\partial t} = \frac{\partial V}{\partial t} \tag{II-67}$$

This equation was derived previously as Eq. II-45.

<u>Example 8</u> - Charged particle in a given electro-magnetic field - Hamiltonian formulation:

We shall use rectangular coordinates x_1, x_2, x_3. The components of the velocity of the particle are v_1, v_2, v_3 and those of the vector potential $\vec{A}$ are A_1, A_2, A_3. The Lagrangian is (cf. Eq. II-56)

$$L \equiv T - V = \sum_{i=1}^{3} \frac{mv_i^2}{2} - q\phi + \frac{q}{c}\vec{A} \cdot \vec{v} \tag{II-68}$$

Here ϕ and A_i may depend on space and time.

As in the previous examples we define the generalized momenta p_i by

$$p_i \equiv \frac{\partial L}{\partial \dot{x}_i} = mv_i + \frac{q}{c} A_i \tag{II-69}$$

Thus, although ordinary rectangular coordinates are used, the generalized momenta are not identical with the ordinary momenta mv_i. This is due to the fact that the potential $\underline{V}$ is velocity dependent. (The definition of momentum given by Eq. II-69 is of great importance in quantum mechanics; see Bohm, Quantum Mechanics, p. 356.

The Hamiltonian $\underline{H}$ is defined by

$$H = \sum_{i=1}^{3} p_i \dot{x}_i - L, \qquad \dot{x}_i = v_i \tag{II-70}$$

One finds

$$H = \sum_i (mv_i + \frac{q}{c} A_i)\, v_i - \sum_i \frac{mv_i^2}{2} + q\phi \frac{q}{c} \quad \vec{A} \cdot \vec{v}$$

$$= \sum_i \frac{mv_i^2}{2} + q\phi = T + q\phi$$

$$= \frac{1}{2m} \sum_i (P_i - \frac{q}{c} A_i)^2 + q\phi \tag{a}$$

Thus, in this case, $\underline{H}$ is $\underline{\text{not}}$ equal to T + V; only the part $q\phi$ of $\underline{V}$ enters in $\underline{H}$. The significance of this can be seen from studying the conservation of $\underline{H}$. We claim without proof (see general derivation below) that the following Hamilton's equations are valid.

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dt} = -\frac{\partial H}{\partial x_i} \tag{II-71}$$

From differential calculus one finds

$$\frac{dH}{dt} = \sum_i \frac{\partial H}{\partial x_i}\frac{dx_i}{dt} + \sum_i \frac{\partial H}{\partial p_i}\frac{dp_i}{dt} + \frac{\partial H}{\partial t} \tag{b}$$

Hamilton's equations and the definition of $\underline{H}$ then give

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} = \frac{\partial V}{\partial t} \tag{II-72}$$

Thus, if ϕ and $\vec{A}$ are independent of time, then the quantity $H = T + q\phi$ is conserved.

The conservation of $T + q \cdot \phi$ may also be seen as follows. (Cf. derivation of conservation laws II-9, II-19.) Newton's equation of motion is

$$m\frac{d\vec{v}}{dt} = \vec{f} = \vec{f}_{el} + \vec{f}_{magn} \tag{II-75}$$

From the definition (II-55c) of $\vec{f}_{magn}$ it is seen that this force is perpendicular to $\vec{v}$ and hence does no work on the particle. Taking the dot product of Eq. II-73 with $\vec{v}$ one finds

$$\vec{v} \cdot m \frac{d\vec{v}}{dt} = \vec{f}_{el} \cdot \vec{v} \tag{c}$$

or

$$\frac{m}{2} \frac{dv^2}{dt} = -q \text{ (grad } \phi) \cdot \vec{v} \tag{d}$$

Since, for $\partial\phi/\partial t = 0$,

$$\frac{d\phi}{dt} = \sum_i \frac{\partial\phi}{\partial x_i} \frac{dx_i}{dt} = (\text{grad } \phi) \cdot \vec{v} \tag{e}$$

one finally derives

$$\frac{d}{dt} \left\{ \frac{mv^2}{2} + q\phi \right\} = 0 \tag{f}$$

or

$$H = T + q\phi = \text{constant} \tag{g}$$

This equation is the same as Eq. II-72 when $\partial H/\partial t = 0$.

5. Hamilton's Equations - General Discussion

We shall now give a general proof of Hamilton's equations, starting from Lagrange's equations. The meaning of the Hamiltonian $\underline{H}$ and the question of its invariance will also be discussed.

We make only the following assumptions: (1) the position of the system is uniquely determined by $\underline{\ell}$ generalized coordinates $q_1 \ldots q_\ell$, (2) there exists a function $L\ (q_1, \ldots q_\ell, \dot{q}_1, \ldots \dot{q}_\ell, t)$, called the Lagrangian, such that the equations of motion can be written in Lagrange's form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0, \qquad i = 1, 2, \cdots \ell \tag{II-43}$$

We then define the generalized (or canonical) momentum corresponding to the coordinate q_i by

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \qquad i = 1, 2, \cdots \ell \tag{II-74}$$

We also define the Hamiltonian by

$$H = \sum_{i=1}^{\ell} p_i \dot{q}_i - L \tag{II-75}$$

In principle, Eqs. II-74 may be solved to give $\dot{q}_i$ as a function of $q_1, \cdots q_\ell, p_1, \cdots p_\ell$ and t. Hence H may be regarded as a function of the coordinates $q_1, \cdots q_\ell$, of their corresponding momenta $p_1, \cdots p_\ell$ and of time t:

$$H = H(q_1, \cdots q_\ell, p_1, \cdots p_\ell, t) \tag{II-76}$$

Thus, by differential calculus, one writes

$$dH = \sum_i \frac{\partial H}{\partial q_i} dq_i + \sum_i \frac{\partial H}{\partial p_i} dp_i + \frac{\partial H}{\partial t} dt \tag{a}$$

On the other hand, Eq. II-75 gives

$$\begin{aligned} dH = {} & \sum_i p_i \, d\dot{q}_i + \sum_i \dot{q}_i \, dp_i \\ & - \sum_i \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \sum_i \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt \end{aligned} \tag{b}$$

From Eq. II-74, it follows that the first and the third terms of the right-hand side of (b) cancel. Furthermore, Lagrange's equations (II-43) may be written

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q_i} \tag{c}$$

(Note that in taking partial derivatives $\underline{L}$ is regarded as a function of the independent variables (q_i), $(\dot{q}_i)$ and $\underline{t}$ whereas $\underline{H}$ is a function of (q_i), (p_i) and $\underline{t}$. Thus in forming $\frac{\partial L}{\partial q_i}$, the (q_j) are kept fixed for all $\underline{j}$ different from $\underline{i}$, furthermore, $\underline{t}$ and all $\dot{q}_j$ are kept fixed.) Hence (b) reduces to

$$dH = \sum_i \dot{q}_i \; dp_i - \sum_i \dot{p}_i \; dq_i - \frac{\partial L}{\partial t} dt \tag{d}$$

Now Eq. II-76 gives the independent variables on which $\underline{H}$ depends. Each of these may be varied without varying the others. Hence, a comparison of (a) and (d) gives Hamilton's equations of motion

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \qquad \frac{dp_i}{dt} = - \frac{\partial H}{\partial q_i} \tag{II-77}$$

and the additional equation

$$\frac{\partial H}{\partial t} = - \frac{\partial L}{\partial t} \tag{II-78}$$

It may be seen that all the Hamiltonian equations for special examples discussed in subsection 4 are special cases of Eqs. II-77. Furthermore, Eqs. II-66, II-72 are special cases of Eq. II-78.

As in subsection 4 one finds

$$\frac{dH}{dt} = \sum_i \frac{\partial H}{\partial q_i}\frac{dq_i}{dt} + \sum_i \frac{\partial H}{\partial p_i}\frac{dp_i}{dt} + \frac{\partial H}{\partial t} \tag{e}$$

Hamilton's equations of motion and Eq. II-78 then give

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t} \tag{II-79}$$

In particular: If L is not a function of time then H is not a function of time either, and Eq. II-79 states that H is invariant, i.e., an integral of the motion. Eq. II-61 is a special case of Eq. II-79.

Eq. II-75 gives a formal definition of H. To study the physical meaning of H we consider the Lagrangian L = T - V. Here T stands for the kinetic energy. Example 8 (charged particle in an electromagnetic field) shows that it is not always justifiable to call V the potential energy. In particular, this example showed that H is not equal to T + V. One should rather consider V as an integral of the forces, i.e., the forces may be derived from V by the differential law (II-53).

Consider now the special case for which the forces are velocity independent. Then V is a function of $q_1, \cdots q_\ell$ and t only, and the generalized forces are given by the simple law

$$Q_j = -\frac{\partial V}{\partial q_j} \tag{II-42}$$

By assumption

$$\frac{\partial V}{\partial \dot{q}_j} = 0 \tag{f}$$

which implies

$$\frac{\partial L}{\partial \dot{q}_j} = \frac{\partial T}{\partial \dot{q}_j} \qquad \text{(g)}$$

Assume, furthermore, that the constraints are independent of time. It then follows from Eq. II-35 that

$$\frac{dx_k}{dt} = \sum_{i=1}^{\ell} \frac{\partial x_k}{\partial q_i} \frac{dq_i}{dt} \qquad \text{(h)}$$

Since, in cartesian coordinates, the kinetic energy is

$$T = \sum_{k=1}^{3n} \frac{m_k}{2} \left(\frac{dx_k}{dt}\right)^2 \qquad \text{(i)}$$

it follows from (h) that its expression in generalized coordinates is of the form

$$T = \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} a_{ij} \frac{dq_i}{dt} \frac{dq_j}{dt} \qquad \text{(II-80)}$$

where the a_{ij} depend on the coordinates q_1, --- q_ℓ only. In other words, T is a homogeneous function of degree two in the $\dot{q}_i$. (The example of the sliding bead shows that this statement is not true, in general, for time-dependent constraints.) Then, by Euler's theorem for homogeneous functions (which may be verified directly from Eq. II-80) and by (g), it follows that

$$\sum_i p_i \dot{q}_i = \sum_i \frac{\partial T}{\partial \dot{q}_i} \dot{q}_i = 2T \qquad \text{(j)}$$

Hence, under the above assumptions,

$$H = 2T - L = T + V \qquad \text{(II-81)}$$

Thus, when the forces are derivable from V, by the simple law (II-42) and when the constraints are time-independent, it is justifiable to call V the potential energy. The Hamiltonian H is then the total energy, $T + V$. This remark applies to Examples 1, 3, and 6 in subsection 4. Furthermore, we note that if in addition L, and hence H, is independent of time then the total energy is conserved, i.e., is an integral of the motion (cf. Examples 1 and 3).

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Aerospace Research Laboratories Applied Mathematics Research Laboratories Wright-Patterson AFB, Ohio 45433	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b GROUP

3. REPORT TITLE

Modern Techniques in Astrodynamics - An Introduction

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Final

5. AUTHOR(S) (First name, middle initial, last name)

Lynn E. Wolaver

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
December 1970	539	-

8a. CONTRACT OR GRANT NO. In-House-Research b. PROJECT NO. 7071-00-22 c. DoD Element 61102F d. DoD Subelement 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) ARL 70-0278

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11 SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
Tech Other	Aerospace Research Laboratories (LB) Wright-Patterson AFB, Ohio 45433

13. ABSTRACT

This report represents lecture notes for a graduate level course in celestial mechanics which has been given at the Air Force Institute of Technology. It covers a review of the two-body problem, discusses the three-body problem, the restricted three-body problem together with regularization and stability analysis. The main portion of the report develops the Hamilton-Jacobi theory and applies it to develop Lagrange's and Gauss' planetary equations. The oblate earth potential is developed and the secular equation solved. Effect of small thrust, drag, lunar-solar gravitational and solar radiation perturbations are developed mathematically and the effects discussed. Von Zeipel's method for the solution of nonlinear equations is developed and used to solve Duffing's equation as an example. Special perturbations are discussed along with errors due to numerical integration and Encke's method is used to obtain approximate analytical results for the motion of stationary satellites. Finally a complete discussion of coordinate systems, time scale and astronomical constants are given. The report ends with a detailed discussion of the shape of the earth. Two appendices briefly cover numerical integration and a review of Lagrangian mechanics.

DD FORM 1 NOV 65 1473

UNCLASSIFIED

Security Classification

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Celestial mechanics Dynamics Trajectories Orbits Lecture Notes von Zeipel Hamilton-Jacobi Applied Mathematics Astrodynamic						

UNCLASSIFIED

Security Classification